Learning to be prepared

Willemien Kets[a] and Mark Voorneveld[a, b,1]

[a] *Dept. of Econometrics and Operations Research, Tilburg University, The Netherlands*

[b] *Dept. of Economics, Stockholm School of Economics, Sweden*

Abstract

Behavioral economics provides several motivations for the common observation that agents appear somewhat unwilling to deviate from recent choices. More recent choices can be more salient than other choices, or more readily available in the agent's mind. Alternatively, agents may have formed habits, use rules of thumb, or lock in on certain modes of behavior as a result of learning by doing. This paper provides discrete-time adjustment processes for strategic games in which players display precisely such a bias towards recent choices. In addition, players choose best replies to beliefs supported by observed play in the recent past, in line with much of the literature on learning. These processes eventually settle down in the minimal prep sets of Voorneveld (2004, 2005).

JEL classification: C72, D83

Keywords: adjustment, learning, minimal prep sets, availability bias, salience, rules of thumb.

SSE/EFI Working Paper Series in Economics and Finance 590,
March 8, 2005

[1]Corresponding author. Address: Department of Economics, Stockholm School of Economics, Box 6501, 113 83 Stockholm, Sweden. E-mail: Mark.Voorneveld@hhs.se. Tel: +46-8-736 92 17. Fax: +46-8-31 32 07.

1. Introduction

The behavioral economics literature provides several motivations for the common observation that agents appear somewhat unwilling to deviate from their recent choices. For instance, Tversky and Kahneman (1982, p. 11) mention the bias towards recent choices as an example of the availability bias, the ease with which instances come to mind. Similarly, Schelling (1960) has argued that players, when indifferent between strategies, choose the most salient strategy. In combination with the recency effect, i.e., the cognitive bias resulting from disproportionate salience of recent stimuli or observations (cf. Miller and Campbell, 1959), this may explain why agents appear to have a preference for recent choices. Other motivations include models for agents displaying defaulting behavior or inertia (cf. Vega-Redondo, 1993, 1995, Madrian and Shea, 2001), the formation of habits (cf. Young, 1998), the use of rules of thumb (cf. Ellison and Fudenberg, 1993), or the locking in on certain modes of behavior due to learning by doing (cf. Grossman et al., 1977) or, as Joosten et al. (1995) express it: unlearning by not doing.

This paper provides a number of discrete-time adjustment processes for mixed extensions of finite strategic games in which players display precisely such a bias towards recent choices. Apart from this behavioral assumption, the assumptions underlying the adaptive processes in this paper are in conformance with much of the literature on learning (cf. Hurkens, 1995, Fudenberg and Levine, 1998, and Young, 1998): players choose best replies to beliefs that are supported by observed play in the recent past. The purpose of this paper is to show that these behaviorally plausible models of adaptive play eventually settle down in so-called minimal prep sets, thus providing a dynamic motivation for such sets.

Minimal prep sets ('prep' is short for 'preparation') were introduced and studied in a static framework in Voorneveld (2004, 2005). This set-valued solution concept for strategic games combines a standard rationality condition, stating that the set of recommended strategies to each player must contain at least one best reply to whatever belief he may have that is consistent with the recommendations to the other players, with players' aim at simplicity, which encourages them to maintain a set of strategies that is as small as possible. This discerns minimal prep sets from *(a)* minimal curb sets (Basu and Weibull, 1991), which are product sets of pure strategies containing not just some, but *all* best responses against beliefs restricted to the recommendations to the remaining players, and *(b)* persistent retracts (Kalai and Samet, 1984), which also require the recommendations to each player to contain at least one best reply to beliefs *in a small neighborhood* of the beliefs restricted to the recommendations to the other players. Voorneveld (2004, 2005) contains a general existence proof and a detailed comparison of minimal prep sets with Nash equilibria, rationalizability, minimal curb sets, and persistent retracts. Tercieux and Voorneveld (2005) show that minimal prep sets provide sharp predictions in many economic applications, including potential games, congestion games, and supermodular games, even in cases where curb sets have no cutting power whatsoever and simply consist of the entire strategy space. This paper thus complements this literature by providing a dynamic motivation for minimal prep sets.

To have play converge to a minimal prep set, players have to somehow 'coordinate' on playing actions from the same minimal prep set. The literature on salience offers some clues on how this coordination process may proceed. Sugden (1995) argues that players need to choose a common probability distribution over their strategies to maximize their expected utility. In our framework, a common pool of probability distributions is provided by the connection between the beliefs of players to which they best-reply and opponents' recent play which induces some coherence in the actions of the players. Moreover, all players always choose the most recent best reply if there are multiple best replies to a given belief. An alternative way of viewing this is provided by Crawford and Haller (1990, p. 577) who speculate that "[..] players begin by searching for a pair of actions to serve as a coordination precedent and then use this precedent to maintain coordination. In effect, they use asymmetric history to "label" actions that cannot be distinguished at the start". Our process does exactly that.

The work that is closest in spirit to our analysis is that of Hurkens (1995). In both his work and in the current paper, convergence to a set-valued solution concept is established, firstly, for discrete-time adjustment processes characterized by conditions on transition probabilities (zero or positive), secondly, for all finite games, (in contrast with e.g. Young (1998), who restricts attention to the special class of weakly acyclic games), and, thirdly, for all memory lengths exceeding a certain lower bound. The main difference between this paper and Hurkens (1995) is that in the latter paper, players choose arbitrary best replies to their beliefs, whereas our players stick to recent best replies.

The outline of this paper is as follows. We recall definitions in Section 2. The evolution of play is discussed in Section 3. Section 4 contains the convergence theorem and explains the steps towards the proof. In Section 5, we discuss our assumptions. Most proofs are contained in the appendix.

2. Preliminaries

Weak set inclusion is denoted by $\subseteq$, strict set inclusion by $\subset$. The number of elements in a finite set S is denoted by $|S|$. For $n \in \mathbb{N}$, the n-fold cartesian product $\times_{i=1}^{n} S$ is denoted by S^n.

A *game* is a tuple $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$, where N is a nonempty, finite set of players, each player $i \in N$ has a nonempty, finite set A_i of pure strategies/actions and a von Neumann-Morgenstern utility function $u_i : A \to \mathbb{R}$ on the set of pure strategy profiles $A = \times_{i \in N} A_i$. Let X_i be a nonempty subset of A_i. The set of mixed strategies of player $i \in N$ with support in X_i is denoted by $\Delta(X_i)$. Payoffs are extended to mixed strategies in the usual way. Let $i \in N$ and let $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(A_j)$ be a belief[2] of player i. The set

$$BR_i(\alpha_{-i}) = \{a_i \in A_i \mid \forall b_i \in A_i : u_i(a_i, \alpha_{-i}) \geq u_i(b_i, \alpha_{-i})\}$$

is the set of pure best responses of player i against α_{-i}.

Fix a game $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$. A *prep set* (Voorneveld, 2004) is a nonempty product set $X = \times_{i \in N} X_i \subseteq A$ of pure-strategy profiles such that for each $i \in N$ and each belief α_{-i} of player i with support in X_{-i}, the set X_i contains *at least one* best response of player i against his belief:

$$\forall i \in N, \forall \alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(X_j) : \quad BR_i(\alpha_{-i}) \cap X_i \neq \emptyset.$$

A prep set X is *minimal* if no prep set is a proper subset of X. Establishing existence of minimal prep sets in finite games is simple: the entire pure-strategy space A is a prep set. Hence the collection of prep sets is nonempty, finite (since A is finite) and partially

[2]Beliefs are thus profiles of mixed strategies: correlation is not allowed.

ordered by set inclusion. Consequently, a minimal prep set exists. See Voorneveld (2004, Thm. 3.2) for a general existence result.

In our adaptive processes, a game $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ is played once every period in discrete time. A *history (of play)* is a sequence $h = (a^1, \dots, a^L) \in A^L$ of some arbitrary length $L \in \mathbb{N}$, whose leftmost element

$$\ell(h) := a^1 \in A$$

is interpreted as the action profile chosen in the previous period according to history h, with $\ell_i(h) := a_i^1 \in A_i$ the action played by $i \in N$. Generally, the k-th element from the left is the action profile $a^k \in A$ chosen $k \in \mathbb{N}$ periods ago.

A *successor* of $h = (a^1, \dots, a^L)$ is a history obtained after one more period of play, a history $h' = (b^1, b^2, \dots, b^{L+1})$ obtained from h by appending a new leftmost element: $b^1 \in A$ and $b^k = a^{k-1}$ for all $k = 2, \dots, L+1$.

Fix a history $h = (a^1, \dots, a^L)$ and a player $i \in N$. The set of actions chosen by i during the previous $k \in \{1, \dots, L\}$ rounds of history h is denoted by

$$p_i(h, k) := \{a_i^1, \dots, a_i^k\}.$$

Assuming that all players' actions were chosen at least once in history h, the *order* $o_{i,h} : \{1, \dots, |A_i|\} \to A_i$ of player i's actions in history h is defined as follows: his first encountered action is $o_{i,h}(1) := a_i^1$ and for $k = 2, \dots, |A_i|$, the k-th encountered action is $o_{i,h}(k) := a_i^m$ with $m = \min\{n = 1, \dots, L \mid a_i^n \notin \{o_{i,h}(1), \dots, o_{i,h}(k-1)\}\}$.

Example 2.1 Consider a two-player game with $N = \{1, 2\}$ and action spaces $A_1 = \{T, B\}, A_2 = \{L, R\}$. Consider the history

$$h = ((T, R), (B, R), (B, L))$$

of length three. Then $\ell(h) = (T, R)$. The set of actions player 1 chose during the most recent two periods is $p_1(h, 2) = \{T, B\}$, whereas $p_2(h, 2) = \{R\}$. As to orders, player 1's action T is encountered first, then B, so $o_{1,h}(1) = T, o_{1,h}(2) = B$. Similarly, $o_{2,h}(1) = R, o_{2,h}(2) = L$. $\triangleleft$

3. Adaptive play

3.1. State space

This section presents a class of Markov chains that models adaptive play with a bias towards choices from the recent past. A game $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ is played once every period in discrete time. In line with much of the literature on learning models (cf. Hurkens, 1995, Fudenberg and Levine, 1998, Young, 1998), players choose, at each moment in time, best replies to beliefs supported by a limited horizon of observed past play of fixed length $T \in \mathbb{N}$.[3]

Consequently, we define the *state space* H to consist of all histories $h = (a^1, \dots, a^L)$ satisfying the following two conditions:

(i) their length is at least T, i.e., $h \in \cup_{K \in \mathbb{N}, K \geq T} A^K$, and

(ii) h is sufficiently "rich", in the sense that all players' actions were chosen at least once before:

$$\forall i \in N, \forall a_i \in A_i, \exists k \in \{1, \dots, L\} : a_i^k = a_i. \tag{1}$$

The latter assures that the 'most recent' best reply to a belief is well-defined, which we use in our behavioral assumption of bias towards recent choices (P2 in Section 3.2). Relaxations of this and other assumptions are discussed in Section 5.

3.2. Transition probabilities

Having defined the set H of states, we proceed to *transition probability functions* $P : H \times H \to [0, 1]$, where $P(h, h')$ is the probability of moving from state $h \in H$ to state $h' \in H$ in one period and $\sum_{h' \in H} P(h, h') = 1$ for all $h \in H$.

A player's beliefs are based on observed play in the past $T \in \mathbb{N}$ periods. That is, for each state $h \in H$, if the sequence of action profiles played in the past T periods is $(a^1, \dots, a^T) \in A^T$, then player i's beliefs are drawn from a probability measure $\mathbb{P}_{(i,(a^1,\dots,a^T))}$ over the set of beliefs (with its standard topology and Borel σ-algebra)

$$\times_{j \in N \setminus \{i\}} \Delta(\{a_j^1, \dots, a_j^T\}) = \times_{j \in N \setminus \{i\}} \Delta(p_j(h, T))$$

[3] Our adjustment processes are defined for a fixed game G and memory length T; to simplify notation, indices G and T are suppressed.

with support in the product set of actions chosen in the previous T periods.

Moreover, given such a belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,T))$, we assume that player i always chooses the most recent best reply to α_{-i}. Players thus have a bias towards recent choices.

Together, the probability distributions $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$, fixing for each player $i \in N$ and account of recent play $(a^1, \ldots, a^T) \in A^T$ the way beliefs are drawn, and the assumption that players are biased towards recent choices, determine the transition probabilities $P(h,h') \in [0,1]$ for each pair of states $(h,h') \in H \times H$.

Let $\mathscr{P}$ denote the class of transition probability functions P achieved in this way, with $P(h,h') > 0$ if and only if states $h, h' \in H$ satisfy conditions P1 and P2 in Fig. 1. Condition P1 is standard for discrete-time processes, stating that between time periods, the game is played once: the process moves from a history h to one of its successors h'. Condition P2 states, first of all, that the process P is a best-reply process: the action $\ell_i(h') \in A_i$ chosen by each player $i \in N$ is a best reply to some belief α_{-i} about the remaining players' behavior based on recent experience, i.e., with support in $\times_{j \in N \setminus \{i\}} \Delta(p_j(h,T))$. In addition, it models the bias towards recent choices: each player $i \in N$ chooses his most recent best reply to his belief α_{-i}.

P1	h' is a successor of $h := (a^1, \ldots, a^L)$;
P2	For each player $i \in N, \ell_i(h')$ is the most recent best reply to some belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,T))$. Formally: $\ell_i(h') = a_i^k$, where $k = \min\{m \in \mathbb{N} \mid BR_i(\alpha_{-i}) \cap \{a_i^1, \ldots, a_i^m\} \neq \emptyset\}$.

Figure 1: For $P \in \mathscr{P}$, $P(h,h') > 0$ iff $h, h' \in H$ satisfy P1 and P2.

Finally, for each $k \in \mathbb{N}$, let $P^k : H \times H \to [0,1]$ denote the k-step transition probabilities of our Markov process with transition probability function $P \in \mathscr{P}$, so that $P^1 = P$ and $P^k = P \circ P^{k-1}$ for all $k > 1$.

4. Convergence and steps towards the proof

This section presents the main result of this paper. Theorem 4.1 states, for each game G and adjustment process in the class $\mathscr{P}$, that if behavior is based on recent experience of sufficient length T, then play will eventually settle down within a minimal prep set. The steps of the proof are briefly explained in this section; the proof itself is contained in the Appendix.

Given a game G and an adjustment process $P \in \mathscr{P}$, we say that the process *eventually settles down* in a minimal prep set if the probability that the process after k steps is in a state $h \in H$ where

- the most recently played action profile lies in a minimal prep set:

$$\ell(h) \in X \text{ for some minimal prep set } X \text{ of } G$$

- all future action profiles remain inside X:

$$\ell(h') \in X \text{ whenever } P^n(h, h') > 0 \text{ for some } n \in \mathbb{N}, h' \in H,$$

converges to one as k goes to infinity.

Theorem 4.1 *Let $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ be a game. Let the horizon $T \in \mathbb{N}$ of recent past play on which beliefs are based satisfy*

$$T \geq \max \left\{ \sum_{i \in N} |A_i| - |N| + 1, 2|A_1|, \ldots, 2|A_n| \right\}, \tag{2}$$

where $n = |N|$. If $P \in \mathscr{P}$, then play eventually settles down in a minimal prep set of G.

In his convergence result, Hurkens (1995, p. 314) uses as a lower bound on memory length the number

$$\underline{K} = \sum_{i \in N} |A_i| - |N| + 1 + \max_{i \in N} |A_i|,$$

which is strictly larger than our bound in (2) under the standard assumption that $|A_i| \geq 2$ for all $i \in N$. He states, however, that his bound is not tight (ibid, p. 313, line 6).

The proof of Theorem 4.1 proceeds in four steps:

Step 1: Let $h_0 \in H$. The process moves with positive probability in $T-1$ steps to a state $h_1 \in H$ where the product set $\times_{i \in N} p_i(h_0, T) \subseteq A$ of actions played in the past T periods is a prep set.

The intuition behind this result is as follows. If, for some state $g \in H$ and some $k \leq T$, the product set $\times_{i \in N} p_i(g, k)$ is a prep set, then players choose with positive probability actions from this prep set for $T-k$ periods in a row. If on the other hand, $\times_{i \in N} p_i(g, k)$ is not a prep set, then there is a nonempty set of players $i \in N$ with a belief $\alpha^*_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(g, k))$ over play in the past k periods to which $p_i(g, k)$ does not contain a best reply. In that case, one can construct a sequence of states $g_1, g_2, \ldots \in H$ with $g_1 = g$, $P(g_k, g_{k+1}) > 0$ for all $k = 1, 2, \ldots$, such that the sequence of product sets $\times_{i \in N} p_i(g_k, k)$ is strictly increasing w.r.t. set inclusion (see lemma A.1 in the Appendix). All these sets are contained in the finite set A, of action profiles which is prep set. Since there are only finitely many actions, the sequence reaches, after a finite number of steps, a state $g_K \in H$ where $\times_{i \in N} p_i(g_K, K)$ is a prep set. From that state onwards, players choose with positive probability actions from the prep set for $T-K$ periods in a row.

Step 2: From state h_1, the process moves with positive probability in a finite number of steps to a state $h_2 \in H$ where $X := \times_{i \in N} p_i(h_2, T)$ is a minimal prep set.[4]

Technically, the proof of this step proceeds as follows. Let $X = \times_{i \in N} X_i \subseteq \times_{i \in N} p_i(h_1, T)$ be a minimal prep set. The proof of this step relies on the fact that one can — under some conditions — perform so-called neighbor switches: from a state $h \in H$, the process moves with positive probability in T steps to a state $h' \in H$ whose horizon of recent past play is identical to the one in h, except that two neighboring actions of some player have changed places (see Lemma A.6). As all permutations of a finite set can be obtained by a chain of such neighbor switches, the process moves with positive probability from state h_1 to a state h' where, for each player $i \in N$, $p_i(h', |X_i|) = X_i$, i.e. the $|X_i|$ most recent actions of each player i are exactly those in his component of the minimal prep set X. Then it is easy to show that the process moves with positive probability to a state h_2

[4]Some intuition for this result is provided by the work of Sugden (1995): as all players share the same behavioral bias ('play the most recent best reply to your current belief'), and as there is some coherence between players' actions (beliefs are based on recent past play), players manage to 'coordinate' on the same minimal prep set.

within a finite number of steps such that $\times_{i \in N} p_i(h', T) = X$ is a minimal prep set.
Step 3: After reaching state h_2, all action profiles that are played with positive probability lie in X, i.e.

$$\forall k \in \mathbb{N}, \forall h \in H : P^k(h_2, h) > 0 \Rightarrow \ell(h) \in X.$$

In h_2, $\times_{i \in N} p_i(h_2, T) = X$ is a minimal prep set, which by definition contains at least one best reply to whatever belief a player may have about other players' choices from X. Hence, by induction, the actions from minimal prep set X will always be fresher in players' recollection of past play than actions outside X, so that to any belief that each player i may have on opponents' play, there is an action in X_i that is the most recent best reply. Hence, from state h_2 onwards, players $i \in N$ only choose actions from X_i.
Step 4: Starting from an arbitrary history h_0, steps 1 and 2 show that there is a positive probability of proceeding to a history h_2 in a finite number of steps, after which play settles down in a minimal prep set, i.e., a positive probability of proceeding to an absorbing set of states in a finitely many steps. Since the initial history was chosen arbitrarily, this eventually happens with probability one, finishing the proof.

5. Discussion of assumptions

5.1. Modifying the assumption on prior play

To guarantee that the most recent best reply to a given belief is well-defined, states $h \in H$ were assumed to be such that all players' actions were chosen at least once before; see (1). This innocuous assumption is the discrete-time analogon of the common assumption in continuous-time dynamics that the process starts away from the boundary, i.e., in a strategy profile having full support. Relaxing this assumption leads to similar results. We discuss two ways to relax this assumption.

Firstly, actions in minimal prep sets are rationalizable (Voorneveld, 2004, Prop. 3.6), so the proof of Theorem 4.1 continues to hold if (1) is replaced by the weaker assumption that all players' rationalizable actions have been chosen at least once before.

Secondly, suppose we allow for the possibility that best replies to some beliefs may not have been played before. This implies that the 'most recent best reply' to a given belief need not exist. To obtain a well-defined process that models a behavioral bias towards recent 'best' actions, one may proceed as follows. Consider an arbitrary state $h = (a^1, \ldots, a^L) \in H'$, where $H' = \cup_{K \in \mathbb{N}, K \geq T} A^K$ is the collection of histories with length L greater than or equal to the lower bound T on memory length, i.e. we drop condition (1) on the state space. As usual, let each player $i \in N$ draw a belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h, T))$ over recent past play from a probability distribution $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$ and assume that i responds by playing the most recent utility maximizing action from the set $\{a_i^1, \ldots, a_i^L\}$ of actions chosen in the past, i.e., the action a_i^k with

$$k = \min \left\{ m \in \{1, \ldots, L\} \mid u_i(a_i^m) = \max_{a_i \in \{a_i^1, \ldots, a_i^L\}} u_i(a_i, \alpha_{-i}) \right\}. \tag{3}$$

As in our initial class $\mathscr{P}$ of processes, the probability distributions $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$ and the bias towards recent choices in (3), determine the transition probabilities $P(h, h') \in [0, 1]$ for each pair of histories $h, h' \in H'$. If the transition probability $P(h, h')$ is positive, then $h, h' \in H$ satisfy conditions P1 and P2' in Fig. 2. Let $\mathscr{P}'$ denote the class of transition probability functions P achieved in this way, with $P(h, h') > 0$ if and only if states $h, h' \in H'$ satisfy conditions P1 and P2' in Fig. 2.

P1	h' is a successor of $h := (a^1, \ldots, a^L)$;
P2'	For each player $i \in N$, $\ell_i(h')$ is the most recent utility maximizing action among the past actions $\{a_i^1, \ldots, a_i^L\}$ to some belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h, T))$. Formally: $\ell_i(h') = a_i^k$, where $k = \min \left\{ m \in \{1, \ldots, L\} \mid u_i(a_i^m) = \max_{a_i \in \{a_i^1, \ldots, a_i^L\}} u_i(a_i, \alpha_{-i}) \right\}$.

Figure 2: For $P \in \mathscr{P}'$, $P(h, h') > 0$ iff $h, h' \in H'$ satisfy P1 and P2'.

Since $\{a_i^1, \ldots, a_i^L\}$ may be a proper subset of A_i, the utility maximizing action from this subset could be a suboptimal reply to the belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(X_j)$ or a best reply that is not contained in a minimal prep set. Consequently, the process need not convergence to a minimal prep set of the underlying game. It does, however, converge to a minimal prep set of a subgame:

Proposition 5.1 *Let $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ be a game and let $T \in \mathbb{N}$. Let $P \in \mathscr{P}'$ and let $h_0 = (b^1, \ldots, b^L) \in H'$ be an initial state. If the horizon $T \in \mathbb{N}$ of recent past play is* *sufficiently large, then play eventually settles down in a minimal prep set of the subgame* $G' = \langle N, (\{b_i^1, \ldots, b_i^L\})_{i\in N}, (u_i)_{i\in N}\rangle$, *where* u_i *is player* i*'s payoff function restricted to* $\times_{i\in N}\{b_i^1, \ldots, b_i^L\}$.

Proof. By P2', given an initial state $h_0 = (b^1, \ldots, b^L) \in H'$, all chosen action profiles in future states by lie in $\times_{i\in N}\{b_i^1, \ldots, b_i^L\} \subseteq A$. The adjustment process with transition matrix $P \in \mathscr{P}'$ for the original game G then reduces to a best-reply process in $\mathscr{P}$ of the subgame G': the full-support condition (1) is satisfied by restricting the strategy space to $\times_{i\in N}\{b_i^1, \ldots, b_i^L\}$ and choosing actions according to P2' coincides with choosing the most recent best reply to their beliefs in the subgame G'. Applying Theorem 4.1 to the subgame G', we find that the process with transition matrix $P \in \mathscr{P}'$ and initial state $h_0 \in H'$ converges to a minimal prep set of the subgame G'. $\square$

5.2. Allowing for other behavioral biases

To show that processes from $\mathscr{P}$ eventually settle down in minimal prep sets, the proof of steps 1 and 2 of Theorem 4.1 (see Appendix) uses that certain transition probabilities are positive to show that the process can move from any initial state $h_0 \in H$ in a finite number of steps to a state $h_2 \in H$ where $\times_{i\in N} p_i(h_2, T)$ is a minimal prep set. The proof of step 3 uses that certain transition probabilities are zero to show that each player — once such a state h_2 is reached — continues to play action profiles from the minimal prep set. We motivated these conditions on the transition probabilities by assuming that players always choose the most recent best reply to a certain belief. However, any class of adjustment process that respects these conditions on the sign will converge to minimal prep sets. Hence, one can easily extend the class of adjustment processes that converge to minimal prep sets.

Consider the more general adjustment process in which, rather than choosing the most recent best reply to beliefs drawn from recent past play, each player $i \in N$ chooses a response according to a probability distribution (mixed strategy)

$$R_{i,h} \in \Delta(A_i)$$

depending on (1) the account $(a^1, \ldots, a^T)$ of recent past play, and (2) the order in which the players' actions appear in h. That is, for each pair of states $h = (a^1, \ldots, a^L), g = (b^1, \ldots, b^K) \in H$:

$$\left.\begin{array}{rcl} (a^1, \ldots, a^T) & = & (b^1, \ldots, b^T) \\ o_{i,h} = o_{i,g} & & \text{for all } i \in N \end{array}\right\} \Rightarrow R_{i,h} = R_{i,g} \text{ for all } i \in N. \tag{4}$$

Example 5.2 In processes from the class $\mathscr{P}$, the probability that $i \in N$ in state $h = (a^1, \ldots, a^L) \in H$ chooses action $a_i \in A_i$ equals the probability of drawing a belief to which a_i is the most recent best reply:

$$R_{i,h}(a_i) = \mathbb{P}_{(i,(a^1,\ldots,a^T))}\left(\{\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,T)) : a_i = a_i^k\}\right),$$

where $k = \min\{m = 1, \ldots, L : BR_i(\alpha_{-i}) \cap \{a_i^1, \ldots, a_i^m\} \neq \emptyset\}$. $\triangleleft$

The collection of functions $R = (R_{i,h})_{i \in N, h \in H}$ determines, for each pair of states $h, h' \in H$, the transition probability $P_R(h, h') \in [0, 1]$. If $P_R(h, h') > 0$, then h' is a successor of h (property P1 in Fig 1) and

$$P_R(h, h') = \prod_{i \in N} R_{i,h}(\ell_i(h)),$$

is the probability of the players choosing action profile $\ell(h')$. Let $\widetilde{\mathscr{P}}$ denote the collection of such transition probability functions $\{P_R : H \times H \to [0, 1] \mid R = (R_{i,h})_{i \in N, h \in H}\}$ with the following two properties: for each pair of histories $h, h' \in H$,

(α) if P1 and P2 hold, then $P_R(h, h') > 0$.

(β) if the product set of actions played during the most recent $k \geq T$ rounds of h is a minimal prep set, play settles down within this set. Formally, if $X := \times_{i \in N} p_i(h, k)$ is a minimal prep set for some $k \geq T$ and $P_R(h, h') > 0$, then $\times_{i \in N} p_i(h', k+1) = X$, i.e., $\ell(h') \in X$.

Proposition 5.3 *Let $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N}\rangle$ be a game and let $T \in \mathbb{N}$. Then $\mathscr{P} \subseteq \widetilde{\mathscr{P}}$. Moreover, if $P_R \in \widetilde{\mathscr{P}}$ and the horizon $T \in \mathbb{N}$ of recent past play is sufficiently large, then play eventually settles down in a minimal prep set of G.*

Proof. Let $P \in \mathscr{P}$. By Example 5.2, there are functions $R = (R_{i,h})_{i \in N, h \in H}$ such that $P = P_R$. Conditions (α) and (β) follow trivially from P1 and P2 in the definition of $\mathscr{P}$. Conclude that $P \in \widetilde{\mathscr{P}}$. The proof of Theorem 4.1 (see Appendix) applies with minor changes to P_R as well:

– condition (α) guarantees that steps 1 and 2 hold without change,

– condition (β) guarantees that step 3 holds without change,

– by (4), there are only finitely many different functions in $R = (R_{i,h})_{i \in N, h \in H}$, so the equivalence relation in step 4 is well-defined and there are again finitely many equivalence classes; hence, also step 4 holds. $\square$

The set inclusion in Proposition 5.3 is strict. One easily finds processes in $\widetilde{\mathscr{P}} \setminus \mathscr{P}$ by letting players choose more freely among recent best replies:

Example 5.4 For $h = (a^1, \ldots, a^L) \in H$ and $i \in N$, let $Y_i(h) \subseteq A_i$ denote the nonempty set of actions which are the most recent best reply to some belief over recent past play:

$$\begin{aligned} a_i \in Y_i(h) \quad \Leftrightarrow \quad & \exists \alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,T)) : a_i = a_i^k, \text{ where} \\ & k = \min\{m = 1, \ldots, L \mid BR_i(\alpha_{-i}) \cap \{a_i^1, \ldots, a_i^m\} \neq \emptyset\}. \end{aligned}$$

To make sure that condition (α) holds, $R_{i,h}$ must assign positive probability to each action in $Y_i(h)$. But player i can choose more freely among recent best replies, not just the *most* recent ones. Let

$$Z_i(h) = BR_i(\times_{j \in N \setminus \{i\}} \Delta(p_j(h,T))) \cap p_i(h,T)$$

be the set of all of i's best replies to beliefs over $\times_{j \in N \setminus \{i\}} p_j(h,T)$ during the horizon of recent past play. Fix a probability distribution over A_i whose support is $Y_i(h) \cup Z_i(h)$. For the purpose of illustration, we take a simple uniform distribution:[5]

$$\forall a_i \in A_i : R_{i,h}(a_i) = \begin{cases} 1/|Y_i(h) \cup Z_i(h)| & \text{if } a_i \in Y_i(h) \cup Z_i(h), \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

[5] Alternatively, one could for instance assign higher probability to the most recent best replies in $Y_i(h)$ than to less recent best replies in $Z_i(h) \setminus Y_i(h)$.

With $R = (R_{i,h})_{i\in N, h\in H}$ as in (5), it follows easily that $P_R \in \widetilde{\mathscr{P}}$. Condition (4) is satisfied because $Y_i(h) = Y_i(g)$ and $Z_i(h) = Z_i(g)$ whenever states $h, g \in H$ satisfy the conditions in (4). Condition (α) is satisfied since each $i \in N$ assigns positive probability to the actions in $Y_i(h)$. Also condition (β) is satisfied: if $X := \times_{i\in N} p_i(h,k)$ is a minimal prep set for some $k \geq T$, then $Y_i(h) \subseteq X_i$ and $Z_i(h) \subseteq p_i(h,T) \subseteq X_i$ for all $i \in N$. Hence, using (5), $\ell_i(h') \in Y_i(h) \cup Z_i(h) \subseteq X_i$ for all $i \in N$, i.e., $\ell(h') \in X$. Finally, since the process also assigns positive probability to possible other recent best replies over observed past play during the last T rounds, $P_R \notin \mathscr{P}$. $\triangleleft$

6. Discussion

The purpose of this paper was to study discrete-time best-response processes with an intuitively appealing bias towards recent actions. Such processes were shown to settle down in minimal prep sets. Several modifications of these processes were discussed in the previous section. There remain, of course, interesting directions for future research, including studying the effect of:

- random perturbations in the processes described above by introducing mistake probabilities or experimentation as in Young (1998),
- introducing players with different levels of sophistication as in Milgrom and Roberts (1991),
- other types of behavioral biases.

Hurkens (1995) already takes up the first two directions in variants of his model. To avoid too much overlap of ideas, we therefore choose not to treat them here. One observation might be useful. Following the discussion in Hurkens (1995, p. 326), one can show that the introduction of perturbations adds little cutting power in two-player games. In contrast with Young's perturbed processes, for instance, this will not lead to a distinction between payoff- or risk-dominant outcomes.

The third direction is the least traditional and therefore the most challenging, but it lies outside the scope of the current paper. We cannot possibly do justice to the long list of choice biases discussed in the behavioral economics literature. Whether other types of biases than the type discussed here give rise to convergence to other solution concepts, is an avenue for further research.

Acknowledgments

We are grateful to the Netherlands Organization for Scientific Research and the Wallander/Hedelius Foundation for partial funding of this research. We thank Wieland Müller and Jörgen Weibull for helpful comments and discussions.

A. Appendix: Proof of Theorem 4.1

Fix a game $G = \langle N, (A_i)_{i\in N}, (u_i)_{i\in N}\rangle$, length $T \in \mathbb{N}$ of recent past play with $T \geq \max\{\sum_{i\in N} |A_i| - |N| + 1, 2|A_1|, \ldots, 2|A_n|\}$, where $n = |N|$, and an adjustment process with transition probability function $P \in \mathscr{P}$. We start with some additional notation. Fix an arbitrary history $h = (a^1, \ldots, a^L) \in H$ and player $i \in N$.

The action player i chose in h a number of $t \in \{1, \ldots, T\}$ periods ago is denoted by

$$a_i(h, t) = a_i^t$$

and the action player i chose in h exactly T periods ago is denoted by

$$\tau_i(h) = a_i^T = a_i(h, T).$$

Action $a_i \in p_i(h, T)$ is *blocked in* h if there is no belief $\alpha_{-i} \in \times_{j\in N\setminus\{i\}}\Delta(p_j(h, T))$ against which it is the most recent best reply.

Finally, player i chose action $a_i \in p_i(h, T)$ during the past T rounds with *frequency*

$$f_i(h, a_i) = |\{t \in \{1, \ldots, T\} : a_i(h, t) = a_i\}|.$$

We now prove the four steps of Theorem 4.1. In Section A.1 step 1 is proven. Sections A.2, A.3, and A.4 provide the building blocks for the proof of step 2, which is proven together with step 3 and 4 in Section A.5.

A.1. Proof of step 1

Step 1: Let $h_0 \in H$. The process moves with positive probability in $T-1$ steps to a state $h_1 \in H$ where the product set $\times_{i \in N} p_i(h_1, T) \subseteq A$ of actions played in the past T periods is a prep set. The proof uses the following lemma:

Lemma A.1 *Consider state $h = (a^1, \ldots, a^L) \in H$ and a number $t \in \{1, \ldots, T-1\}$.*

(a) Suppose that $\times_{i \in N} p_i(h,t) \subseteq A$ is not a prep set. Then the process moves with positive probability to a successor $h' \in H$ where

$$\times_{i \in N} p_i(h,t) \subset \times_{i \in N} p_i(h', t+1). \tag{6}$$

(b) Suppose that $\times_{i \in N} p_i(h,t) \subseteq A$ is a prep set. Then the process moves with positive probability to a successor $h' \in H$ where

$$\times_{i \in N} p_i(h,t) = \times_{i \in N} p_i(h', t+1). \tag{7}$$

Proof. (a): Since $\times_{i \in N} p_i(h,t) \subseteq A$ is not a prep set, there is a nonempty set $S \subseteq N$ of players $i \in N$ with a belief $\alpha^*_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,t))$ over the play in the past t periods to which $p_i(h,t)$ does not contain a best reply: $BR_i(\alpha^*_{-i}) \cap p_i(h,t) = \emptyset$. Fix such a belief α^*_{-i} for each $i \in S$ and let $b_i \in A_i$ be the most recently played best reply to α^*_{-i} in h:

$$b_i = a_i^k, \text{ where } k = \min\{m \in \mathbb{N} \mid BR_i(\alpha^*_{-i}) \cap \{a_i^1, \ldots, a_i^m\} \neq \emptyset\}.$$

For each $i \in N \setminus S$, let $b_i \in p_i(h,t)$ be the most recent best reply to an arbitrary belief over play in the past t periods. Such a best reply exists by definition of S. By P1 and P2, the process moves with positive probability from state h to successor $h' = (b, a^1, \ldots, a^L)$. Now (6) holds by construction: if $i \in N \setminus S$, then $b_i \in p_i(h,t)$, so $p_i(h,t) = p_i(h', t+1)$, and if $i \in S$, then $b_i \notin p_i(h,t)$, so $p_i(h,t) \subset p_i(h,t) \cup \{b_i\} = p_i(h', t+1)$.

(b): Fix, for each $i \in N$, a belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,t))$ over the play in the past t periods. Since $\times_{i \in N} p_i(h,t)$ is a prep set, there is an action $b_i \in p_i(h,t)$ which is the most recent best reply to this belief. By P1 and P2, the process moves with positive probability from h to $h' = (b, a^1, \ldots, a^L)$. Since $b_i \in p_i(h,t)$ for all $i \in N$, it follows that

$p_i(h', t+1) = p_i(h, t)$, so (7) holds. $\square$

Applying Lemma A.1 $T-1$ times, one can construct a sequence $g_1, \ldots, g_T$ in H with $g_1 := h_0$ and for all $k = 1, \ldots, T-1$: $P(g_k, g_{k+1}) > 0$ and

$$\times_{i\in N} p_i(g_k, k) \subseteq \times_{i\in N} p_i(g_{k+1}, k+1),$$

with strict inclusion if $\times_{i\in N} p_i(g_k, k)$ is not a prep set and equality otherwise. The sequence of product sets $\times_{i\in N} p_i(g_k, k)$ in A can increase strictly during at most $\sum_{i\in N} |A_i| - |N|$ steps: the action space A is a prep set containing $\sum_{i\in N} |A_i|$ actions; $\times_{i\in N} p_i(g_1, 1)$ captures $|N|$ of them, and in each step at least one action is added until a prep set is reached. Hence, the sequence has to reach, after $K \leq \sum_{i\in N} |A_i| - |N|$ steps, a state $g_{K+1} \in H$ where $\times_{i\in N} p_i(g_{K+1}, K+1)$ is a prep set[6]. In the final $T-K-1$ steps, we proceed to a state g_T, where

$$\times_{i\in N} p_i(g_T, T) = \times_{i\in N} p_i(g_{T-1}, T-1) = \cdots = \times_{i\in N} p_i(g_{K+1}, K+1)$$

remains a prep set. Taking $h_1 := g_T$ finishes the proof of step 1.

A.2. States without blocked actions

In this section, we show that from a state $h \in H$ such that $\times_{i\in N} p_i(h, T)$ is a prep set, the process moves with positive probability within a finite number of steps to a state $h' \in H$ where $\times_{i\in N} p_i(h', T) \subseteq \times_{i\in N} p_i(h, T)$ is a prep set without blocked actions. This is established in Lemma A.3, using Lemma A.2. Furthermore, in Remark A.4 we show that when considering a sequence $g_1, \ldots, g_K$ such that, for all $k = 1, \ldots, K$, $\times_{i\in N} p_i(g_k, T)$ is a prep set and $\times_{i\in N} p_i(g_1, T) \supseteq \cdots \supseteq \times_{i\in N} p_i(g_K, T)$, we can assume without loss of generality that none of the states $(g_k)_{k=1,\ldots,K}$ contains a blocked action. We use this result in the lemmata discussed in the next subsections.

Lemma A.2 *Let $h \in H$ be such that $\times_{i\in N} p_i(h, T)$ is a prep set. For each player $i \in N$, define $\beta_i(h) \in p_i(h, T)$ as follows:*

[6] This motivates the lower bound $L := \sum_{i\in N} |A_i| - |N| + 1$ on T in (2): reaching a prep set can take $L-1$ steps; recalling the added actions and those in g_1 can consequently take a memory length $L+1$. The same reasoning applies to other adjustment processes in the literature; cf. Hurkens (1995, p. 314).

- *if $\tau_i(h)$ is blocked, let $\beta_i(h) \in p_i(h,T)$ be an arbitrary non-blocked action.*

- *if $\tau_i(h)$ is not blocked, let $\beta_i(h) = \tau_i(h)$.*

Set $h' = (\beta(h); h)$, with $\beta(h) = (\beta_i(h))_{i \in N}$. Then:

$$
\begin{aligned}
P(h,h') &> 0 && (8)\\
\times_{i \in N} p_i(h',T) &\subseteq \times_{i \in N} p_i(h,T) && (9)\\
\times_{i \in N} p_i(h',T) & \quad \textit{is a prep set.} && (10)
\end{aligned}
$$

Proof. For all $i \in N$, $\beta_i(h) \in p_i(h,T)$ is not blocked by definition: there is a belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h,T))$ against which $\beta_i(h)$ is the most recent best reply. By P1 and P2, (8) holds. Since $\beta_i(h) \in p_i(h,T)$ for all $i \in N$, (9) holds. To prove (10), let $i \in N$ and $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(p_j(h',T))$. To show: $BR_i(\alpha_{-i}) \cap p_i(h',T) \neq \emptyset$. By construction, $p_i(h',T)$ equals either $p_i(h,T)$ or, if $\tau_i(h)$ was blocked and chosen only once in the most recent T periods of history h, $p_i(h,T) \setminus \{\tau_i(h)\}$. Consequently, $p_i(h',T)$ still contains a best reply to every belief over $\times_{j \in N \setminus \{i\}} \Delta(p_j(h,T))$, in particular to every belief over the subset $\times_{j \in N \setminus \{i\}} \Delta(p_j(h',T))$. □

Claim (9) means that we weakly decrease the pool of feasible beliefs in going from h to $h' = (\beta(h); h)$. This implies that if $a_i := \tau_i(h)$ was blocked in h, but was chosen more than once in the last T rounds of h, i.e., if $a_i \in p_i(h',T)$, then it remains blocked:

$$
\text{if } a_i := \tau_i(h) \text{ was blocked in } h \text{ and } a_i \in p_i(h',T), \text{ then it is blocked in } h'. \tag{11}
$$

By definition, blocked actions are not chosen in going from h to h'. Thus, if an action is blocked in h, it is either no longer contained in $\times_{i \in N} p_i(h',T)$, in which case (9) holds with strict inclusion, or it remains blocked in h' by (11), but lies further back in players' memory. Hence, repeated application of Lemma A.2 to the sequence $g_1, g_2, \ldots$ in H with $g_1 = h$ and for all $k \in \mathbb{N} : g_{k+1} = (\beta(g_k); g_k)$, yields that a blocked action disappears from memory in at most T steps, in which case the product set of recent actions has become strictly smaller in the weakly decreasing sequence

$$
\times_{i \in N} p_i(g_1,T) \supseteq \times_{i \in N} p_i(g_2,T) \supseteq \cdots
$$

By (10), the product set remains a prep set. Since there are only finitely many prep sets, it follows that we eventually reach a state g_k without blocked actions. This proves:

Lemma A.3 *Let $h \in H$ be such that $\times_{i\in N} p_i(h,T)$ is a prep set. Either h contains no blocked actions, or the process moves with positive probability in a finite number of steps to a state $h' \in H$ where $\times_{i\in N} p_i(h',T) \subset \times_{i\in N} p_i(h,T)$ is a prep set and h' contains no blocked actions.*

Remark A.4 As part of the proof of step 2, we will establish the following: Given a state $g_1 \in H$ where $\times_{i\in N} p_i(g_1,T)$ is a prep set, the process moves with positive probability in a finite number of steps to a state $g_K \in H$ where $\times_{i\in N} p_i(g_K,T)$ is a prep set with the property that there is a minimal prep set $X = \times_{i\in N} X_i \subseteq \times_{i\in N} p_i(g_1,T)$ such that for each $i \in N$:

$$p_i(g_K, |X_i|) = X_i, \tag{12}$$

i.e., for each player $i \in N$, the most recent $|X_i|$ actions are exactly those in i's component of the minimal prep set X. We do so by showing that with positive probability the process moves through a sequence of states $g_1, g_2, \ldots, g_K$ such that for all $k = 1, \ldots, K-1$: $\times_{i\in N} p_i(g_k,T)$ is a prep set and

$$\times_{i\in N} p_i(g_1,T) \supseteq \times_{i\in N} p_i(g_2,T) \supseteq \cdots \supseteq \times_{i\in N} p_i(g_K,T). \tag{13}$$

If any of these states g_k contains a blocked action, we can apply Lemma A.3 to move to a state g' where $\times_{i\in N} p_i(g',T) \subset \times_{i\in N} p_i(g_k,T)$ is a prep set and g' contains no blocked actions. Then, we can resume the process to arrive at a prep set as in (12) from g'. Since there are only finitely many prep sets and the prep set $\times_{i\in N} p_i(g',T)$ is strictly contained in $\times_{i\in N} p_i(g_k,T)$, we eventually reach — in a finite number of steps — a state from which we can apply the process to reach a state g_K with the property as in (12) without ever encountering a state in which one of the actions is blocked. Hence: *we can assume without loss of generality that none of the states $(g_k)_{k=1,\ldots,K}$ contains a blocked action.* $\triangleleft$

A.3. Drag-to-front operations and frequency correction

Consider a state $h \in H$ containing no blocked actions for which $\times_{i\in N} p_i(h,T)$ is a prep set. Then, by definition (see Lemma A.2), for each $i \in N$, $\beta_i(h) = \tau_i(h)$, the action i

chose T periods ago. Hence, in the successor $(\beta(h); h) = (\tau(h); h)$, this action is dragged to the front of player i's account of recent past play. For easy reference, call the transition from h to $(\beta(h); h) = (\tau(h); h)$ a *drag-to-front operation.*

Suppose some player $j \in N$ has an action $a_j \in p_j(h, T)$ with frequency $f_j(h, a_j) = 1$. Since $T \geq 2|A_j|$ by (2), there must be an action $b_j \in p_j(h, T)$ with frequency $f_j(h, b_j) \geq 3$. By Remark A.4, and using drag-to-front-operations if necessary, we can assume without loss of generality that player j chose b_j exactly T periods ago: $\tau_j(h) = b_j$. For each player $i \in N$, define $\gamma_i(h') \in p_i(h', T)$ as follows:

$$\gamma_i(h) = \begin{cases} \tau_i(h) & \text{if } i \neq j, \\ a_j & \text{if } i = j. \end{cases} \tag{14}$$

Set $h' = (\gamma(h); h)$ with $\gamma(h) = (\gamma_i(h))_{i \in N}$. Recall: (1) $\gamma_i(h) \in p_i(h, T)$ for all $i \in N$, (2) $\times_{i \in N} p_i(h, T)$ is a prep set, and (3) no actions in h are blocked; so each $\gamma_i(h)$ is the most recent best reply to a belief $\alpha_{-i} \in \times_{k \in N \setminus \{i\}} \Delta(p_k(h, T))$. By P1 and P2, $P(h, h') > 0$.

By construction, $\times_{i \in N} p_i(h', T) = \times_{i \in N} p_i(h, T)$ remains a prep set. The frequency of the actions of players $i \neq j$ is unaffected: $\forall i \in N \setminus \{j\}, \forall c_i \in p_i(h', T) = p_i(h, T)$: $f_i(h', c_i) = f_i(h, c_i)$. For player j and $c_j \in p_j(h', T) = p_j(h, T)$:

$$f_j(h', c_j) = \begin{cases} f_j(h, c_j) & \text{if } c_j \notin \{a_j, b_j\}, \\ f_j(h, a_j) + 1 = 2 & \text{if } c_j = a_j, \\ f_i(h, b_j) - 1 \geq 2 & \text{if } c_j = b_j. \end{cases}$$

By going from h to h', the number of actions with frequency one has strictly decreased, whereas there is no action with frequency larger than or equal to two whose frequency becomes less than two.

Repeating this process, we eventually reach a state where all actions in the history of recent past play have frequency greater than or equal to two. By Lemma A.3, we may assume that none of its actions is blocked. This proves:

Lemma A.5 *Let $h \in H$ be such that $\times_{i \in N} p_i(h, T)$ is a prep set. Then the process moves with positive probability in a finite number of steps to a state $h' \in H$ with $\times_{i \in N} p_i(h', T) \subseteq \times_{i \in N} p_i(h, T)$ such that*

[C1] $\times_{i \in N} p_i(h', T)$ *is a prep set,*

[C2] *all actions have frequency at least 2:* $\forall i \in N, \forall a_i \in p_i(h', T) : f_i(h', a_i) \geq 2$,

[C3] h' *contains no blocked actions.*

A.4. Neighbor switches

If we repeatedly apply drag-to-front operations starting in a state $h \in H$ such that no actions are blocked and $\times_{i \in N} p_i(h, T)$ is a prep set, we get a sequence of states $g_0, g_1, \dots \in H$ with $g_0 := h$ such that for all players $i \in N$ and all $t \in \mathbb{N}$: $\ell_i(g_t) = \tau_i(g_{t-1})$, i.e., we get a periodic repetition of each player's actions.

Suppose instead that we wish to change the order[7] in which player i plays two neighboring actions, while the others continue to play actions in their given order. For instance, we may want to go from Fig. 3.a to Fig. 3.e, where we switched player i's actions **b** and **c**, chosen 2 and 3 periods ago, respectively, keeping the order of actions of players $j \neq i$ the same. In Fig. 3, the length of recent past play T is 4; actions chosen during the most recent four periods are contained in the boxed part of the table; actions outside the boxes have disappeared from recent past play. For instance, in Fig. 3.c, player i chose c five periods ago, d six periods ago. Since $T = 4$, these actions are no longer part of recent past play.

The idea is simple: we can use drag-to-front operations until the actions we wish to switch are those chosen $T-1$ and T periods ago (the transition from Fig. 3.a to Fig. 3.b); in the next two periods, let players $j \neq i$ continue with drag-to-front operations, while player i chooses the actions we wish to switch in reverse order (in going from Fig. 3.b to Fig. 3.c, i chooses **b** instead of **c**, in going from the Fig. 3.c to Fig. 3.d, i chooses **c** instead of **b**). Finally, we can use drag-to-front operations until the switched actions are back at time slots t and $t+1$ in the recent past play (the transition from Fig. 3.d to Fig. 3.e). Formally:

Lemma A.6 *Let* $h \in H$ *satisfy* [C1] *to* [C3]*. Let* $i \in N, t \in \{1, \dots, T-1\}$*. Assuming w.l.o.g. (Remark A.4) that we encounter no blocked actions, the process moves with*

[7] Although they may be related (for instance in the case of drag-to-front operations), the order in which a player $i \in N$ plays his actions is different from the way in which his actions are encountered in a given history h, i.e., the order $o_{i,h}$ defined in Section 2.

Fig. 3.a	player i:	a	**b**	**c**	d				
	player j:	α	β	γ	δ				

Fig. 3.b	player i:	d	a	**b**	**c**	d			
	player j:	δ	α	β	γ	δ			

Fig. 3.c	player i:	**b**	d	a	**b**	c	d		
	player j:	γ	δ	α	β	γ	δ		

Fig. 3.d	player i:	**c**	**b**	d	a	b	c	d	
	player j:	β	γ	δ	α	β	γ	δ	

Fig. 3.e	player i:	a	**c**	**b**	d	a	b	c	d
	player j:	α	β	γ	δ	α	β	γ	δ

Figure 3: Switch i's actions **a** and **b**, keeping those of players $j \neq i$ in the same order.

positive probability in T steps to a state $h' \in H$ satisfying [C1] *to* [C3] *and in which $a_j(h',k) = a_j(h,k)$ if $j \neq i$ or if ($j = i$ and $k \notin \{t, t+1\}$), whereas $a_i(h',t) = a_i(h,t+1)$ and $a_i(h',t+1) = a_i(h,t)$.*

Proof. For notational convenience, let a_i and b_i be the actions player i chose $t+1$ and t periods ago in h, respectively. Performing $T-t-1$ drag-to-front operations, we reach a state h_1 satisfying [C1] to [C3] in which a_i is the action i chose T periods ago and b_i the action he chose $T-1$ periods ago.

Construct a successor h_2 of h_1 as follows: for each $j \in N \setminus \{i\}$, set $s_j^1 = \tau_j(h_1)$ and set $s_i^1 = b_i$. Define $h_2 = (s^1; h_1)$, where $s^1 = (s_j^1)_{j \in N}$.

Construct a successor h_3 of h_2 as follows: for each $j \in N \setminus \{i\}$, set $s_j^2 = \tau_j(h_2)$ and set $s_i^2 = a_i$. Define $h_2 = (s^2; h_2)$, where $s^2 = (s_j^2)_{j \in N}$.

For players $j \neq i$, these two steps involve simple drag-to-front operations. For player i it involves reversing the order: in going from h_1 to h_2, i chooses b_i, in going from h_2 to h_3, i chooses a_i, rather than playing first a_i, then b_i.

As $\times_{i \in N} p_i(h_1, T)$ is a prep set and no actions are blocked in h_1, it follows from P1 and P2 that $P(h_1, h_2) > 0$. Moreover, as all actions in h have frequency at least 2, we have that $p_i(h_1, T) = p_i(h_2, T)$ for all $i \in N$. Hence, also $\times_{i \in N} p_i(h_2, T)$ is a prep set. By Remark A.4 we may assume that h_2 contains no blocked actions. Hence, also $P(h_2, h_3) > 0$. Moreover, it is easy to see that frequencies in h_3 are identical to frequencies in h_1, i.e., at least equal to 2. We can thus conclude that also h_3 satisfies [C1] to [C3].

In h_3, the two actions that are played most recently are a_i and b_i, respectively. Thus, performing $t - 1$ drag-to-front operations leads to the desired state h'. □

A.5. Proof of the remaining steps

Step 2: Let $h_1 \in H$ be such that $\times_{i \in N} p_i(h_1, T)$ is a prep set. The process moves with positive probability in a finite number of steps to a state $h_2 \in H$ where $\times_{i \in N} p_i(h_2, T)$ is a minimal prep set.

Proof. By Lemma A.5, the process moves with positive probability in a finite number of steps from h_1 to a state $g \in H$ satisfying [C1] to [C3]. Let $X = \times_{i \in N} X_i \subseteq \times_{i \in N} p_i(g, T)$ be a minimal prep set. Assuming w.l.o.g. (Remark A.4) that from g onward we do not encounter blocked actions, Lemma A.6 allows us to perform neighbor switches. Every permutation of a finite set can be obtained by a chain of neighbor switches; thus, repeated application of Lemma A.6 yields that the process moves in a finite number of steps to a state $g_0 \in H$ with the property that for each player $i \in N$, $p_i(g_0, |X_i|) = X_i$, i.e., for each player $i \in N$, the most recent $|X_i|$ actions in g_0 are exactly those in i's component of the minimal prep set X.

For each $k \in \mathbb{N}$, let $g_k := ((a(g_{k-1}, |X_i|))_{i \in N}; g_{k-1}) \in H$, i.e., g_k is the successor of g_{k-1} obtained by letting each player $i \in N$ play the action he chose $|X_i|$ periods ago in g_{k-1}. Recalling that X is a minimal prep set, a simple inductive proof establishes that for all $k \in \mathbb{N}$ it holds that $P(g_{k-1}, g_k) > 0$ and for all players $i \in N$ we have

$$p_i(g_k, \min\{|X_i| + k, T\}) = X_i.$$

Set $k = T$ to find that $\times_{i \in N} p_i(h_T, T) = X$. Taking $h_2 := g_T$ finishes the proof of step 2. □

Step 3: Let $h_2 \in H$ be such that $X = \times_{i \in N} p_i(h_2, T)$ is a minimal prep set. After reaching h_2, all action profiles that are played with positive probability lie in X:

$$\forall k \in \mathbb{N}, \forall h \in H : P^k(h_2, h) > 0 \Rightarrow \ell(h) \in X. \tag{15}$$

Proof. By P1 and P2, players always base beliefs on the actions played in the last T periods and choose the most recent best reply to such beliefs. In h_2, their account of recent play $\times_{i \in N} p_i(h_2, T)$ equals the minimal prep set X, which by definition contains at least one best reply to whatever belief a player may have about other players' choices from X. Hence, by induction, the actions from minimal prep set X will always be fresher in players' recollection of past play than actions outside X, i.e., beliefs and best replies to these beliefs will, by P1 and P2, always have support in X. Formally, for all $k \in \mathbb{N}$ and $h \in H$:

$$\text{if } P^k(h_2, h) > 0, \text{ then } \times_{i \in N} p_i(h, T + k) = X,$$

and hence

$$\times_{i \in N} p_i(h, T) \subseteq X.$$

In particular, this means $\ell(h) \in X$, i.e., (15) holds. □

Step 4: For every state $h_0 \in H$, the process eventually reaches a state $h_2 \in H$ satisfying the conditions in step 2, i.e., where according to step 3 play settles down in a minimal prep set.

Proof. Recall from Section 3.2 that in each state $h \in H$, if the sequence of action profiles from the past T periods is $(a^1, \ldots, a^T) \in A^T$, then player i's beliefs are drawn from a probability distribution $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$. By P2, he chooses the most recent best reply to every such belief. Thus, player i's choice behavior is the same in two states $h = (a^1, \ldots, a^L), g = (b^1, \ldots, b^K) \in H$ if they have the same account of recent past play, i.e. if $(a^1, \ldots, a^T) = (b^1, \ldots, b^T)$, and the same the order in which player i's actions are encountered, i.e., if $o_{i,g} = o_{i,h}$. Since there are only finitely many elements in A^T and N, it follows that the set of positive transition probabilities $\{P(h, h') \mid P(h, h') > 0\}$ is a finite set. Let $\varepsilon > 0$ be its minimum.

By steps 1 to 3, it is possible, from any history $h_0 \in H$, to reach a state $h_2 \in H$ in an absorbing set where play settles down in a minimal prep set in a finite number of steps, say $k(h_0) \in \mathbb{N}$. By definition of equivalence, $k(h) = k(h_0)$ for all $h \in [h_0]$: the set $\{k(h_0) \mid h_0 \in H\}$ is finite. Let $\kappa \in \mathbb{N}$ be its minimum.

By definition of ε and κ, the probability of entering an absorbing set where play settles down in a minimal prep set in at most κ steps is at least ε^κ from any state. Hence, the probability of not reaching an absorbing set in κ steps is at most $1-\varepsilon^\kappa$, which is less than 1. So the probability of not reaching an absorbing set in $k\kappa$ steps is less than or equal to $(1-\varepsilon^\kappa)^k$, which goes to zero as k goes to infinity. $\square$

References

Basu, K., Weibull, J.W., 1991. Strategy subsets closed under rational behavior. Econ. Letters 36, 141-146.

Crawford, V. P. and Haller, H., 1990. Learning how to cooperate: optimal play in repeated coordination games. Econometrica 58, 571-595.

Ellison, G., Fudenberg, D., 1993. Rules of thumb for social learning. J. Political Economy 101, 612-643.

Fudenberg, D., Levine, D.K., 1998. The theory of learning in games. Cambridge MA: MIT Press.

Grossman, S.J., Kihlstrom, R.E., Mirman, L.J., 1977. A Bayesian approach to the production of information and learning by doing. Rev. Econ. Stud. 44, 533-547.

Hurkens, S., 1995. Learning by forgetful players. Games Econ. Behav. 11, 304-329.

Joosten, R., Peters, H., Thuijsman, F., 1995. Unlearning by not doing: Repeated games with vanishing actions. Games Econ. Behav. 9, 1-7.

Kalai, E., Samet, D., 1984. Persistent equilibria in strategic games. Int. J. Game Theory 14, 41-50.

Madrian, B.C., Shea, D.F., 2001. The power of suggestion: Inertia in 401(k) participation and savings behavior. Quart. J. Econ. 116, 1149-1187.

Milgrom, P., Roberts, J., 1991. Adaptive and sophisticated learning in normal form games. Games Econ. Behav. 3, 82-100.

Miller, N., Campbell, D. T., 1959. Recency and primacy in persuasion as a function of the timing of speeches and measurements. J. Abnormal and Social Psychology 59, 1-9.

Schelling, T., 1960. The strategy of onflict. Cambridge MA: Harvard University Press.

Sugden, R., 1995. A theory of focal points. Econ. J. 105, 533-550.

Tercieux, O., Voorneveld, M., 2005. The cutting power of preparation. SSE/EFI working paper series in economics and finance, No. 583, Stockholm School of Economics, Sweden.

Tversky, A., Kahneman, D., 1982. Judgment under uncertainty: Heuristics and biases. In: Kahneman, D., Slovic, P., and Tversky, A. (eds.). Judgment under uncertainty: heuristics and biases, Cambridge: Cambridge University Press, pp. 3-20.

Vega-Redondo, F., 1993. Simple and inertial behavior: an optimizing decision model with imprecise perceptions. Econ. Theory 3, 87-98.

Vega-Redondo, F., 1995. The dynamics and revealed preference of status-quo optimality. Econ. Theory 6, 263-282.

Voorneveld, M., 2004. Preparation. Games Econ. Behav. 48, 403-414.

Voorneveld, M., 2005. Persistent retracts and preparation. Games Econ. Behav. 51, 228-232.

Young, H.P., 1998. Individual strategy and social structure. Princeton NJ: Princeton University Press.