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, or use rules of thumb. This paper provides discrete-time adjustment processes for strategic games in which players display such a bias towards recent choices. In addition, players choose best replies to beliefs supported by observed play in the recent past. We characterize the limit behavior of these processes by showing that they eventually settle down in minimal prep sets [Voorneveld, Games Econ. Behav. 48 (2004) 403 – 414].

JEL classification: C72, D83

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

[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 so-called recency effect (Miller and Campbell, 1959) this may explain why agents appear to have a preference for recent choices. The recency effect refers to the cognitive bias that results from disproportionate salience of recent stimuli or observations. 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 class 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. The latter feature 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.

The choice of the term "preparation" in connection with minimal prep sets is motivated by the rationality requirement. Given an arbitrary belief of a player that is consistent with the recommendations to the other players, his recommended set of strategies leaves him well-prepared: it contains an optimal response against all such eventualities. On the other hand, one does not have to be exhaustive to be prepared: the notion of prep sets avoids the potential snowball effect from the requirement that all best replies against a given belief (and all best replies against all these best replies, and so on...) need to be included, as demanded of the curb sets of Basu and Weibull (1991). Think of the set of recommendations to a player in a minimal prep set as a well-packed suitcase for a holiday: you want to be prepared for different kinds of weather, but bringing all five of your umbrellas and all seven bathing suits may be overdoing it.[2]

The game in Figure 1 provides a simple example to illustrate the difference between pure Nash equilibria, minimal curb sets, and minimal prep sets. The game has no pure Nash equilibria. Its only — hence minimal — curb set is the entire pure strategy space $\{R_1, R_2, R_3\} \times \{C_1, C_2, C_3\}$. There are two minimal prep sets, $\{R_1, R_2\} \times \{C_1, C_2\}$ and $\{R_2, R_3\} \times \{C_2, C_3\}$, roughly speaking the "Matching pennies" subgames.

	C_1	C_2	C_3
R_1	$1, -1$	$-1, 1$	$-100, -100$
R_2	$-1, 1$	$1, -1$	$-1, 1$
R_3	$-100, -100$	$-1, 1$	$1, -1$

Figure 1: A 3×3 game

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. Voorneveld et al. (2005) provide axiomatizations of minimal prep sets and minimal curb sets. 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 minimal curb sets have no cutting power whatsoever and simply consist of the entire strategy space. The current paper complements this literature by providing a dynamic motivation for minimal prep sets.

[2]We are grateful to Dries Vermeulen for suggesting this "no excess luggage" interpretation.

For play to settle down in a specific set, like a minimal prep or curb set, players somehow need to learn to coordinate on actions from within this set. Crawford and Haller (1990, p. 577) indicate that an important coordination device is the fact that players "use asymmetric history to "label" actions that cannot be distinguished at the start". In our model, players use histories to form beliefs, but also to select between best replies. The behavioral bias towards recent best replies is a specific example of a way to exploit the history to distinguish between actions that all perform well against a player's beliefs.

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 weakly acyclic games), and, thirdly, for all memory lengths exceeding a certain lower bound. There are, however, important differences. The behavioral bias towards recent choices that players use to distinguish between best replies is absent in Hurkens' model: there, players indiscriminately choose best replies to their beliefs. As a consequence, players in our model need to keep track of whether one best reply was chosen more recently than another. However, this does not mean that a player needs to have perfect memory of his own past action choices. This is particularly clear if a player has only two actions: if both happen to be a best reply to his current belief, the action he chose in the previous round is the most recent one and therefore all he needs to recall. We return to this issue in more detail in Remarks 3.1 and 5.2.

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 a more general class of adjustment processes for which play also settles down in minimal prep sets. Section 6 contains concluding remarks. All 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 $k \in \mathbb{N}$, the k-fold cartesian product $\times_{i=1}^{k} S$ is denoted by S^k.

A ***game*** is a tuple $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$, where $N = \{1, \ldots, n\}$ 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[3] 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 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 $h = (a^1, \ldots, a^L) \in A^L$ 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 history $h = (a^1, \ldots, a^L)$ is a history obtained after one more period of play, a history $h' = (b^1, b^2, \ldots, b^{L+1})$ obtained from h by appending a new leftmost element: $b^1 \in A$ and $b^k = a^{k-1}$ for all $k = 2, \ldots, L+1$.

Fix a history $h = (a^1, \ldots, a^L)$ and a player $i \in N$. The set of actions chosen by i during the last[4] $k \in \{1, \ldots, L\}$ rounds of history h is denoted by

$$\lambda_i(h, k) := \{a_i^1, \ldots, a_i^k\}.$$

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

[4]Hence our choice of the alliterative λ (lambda).

The ***order*** $o_{i,h}$ of player i's actions in history h is defined as follows: his most recent action, i.e., the first encountered action is $o_{i,h}(1) := a_i^1$ and, inductively, for $k = 2, \ldots, |\{a_i^1, \ldots, a_i^L\}|$, the k-th encountered action is $o_{i,h}(k) := a_i^m$ with $m = \min\{q \in \{1, \ldots, L\} \mid a_i^q \notin \{o_{i,h}(1), \ldots, o_{i,h}(k-1)\}\}$.

Example 2.1 Consider a two-player game with $N = \{1, 2\}$ and action spaces $A_1 = \{R_1, R_2\}, A_2 = \{C_1, C_2\}$. Consider the history

$$h = ((R_1, C_2), (R_2, C_2), (R_2, C_1))$$

of length three. Then $\ell(h) = (R_1, C_2)$. The set of actions player 1 chose during the most recent two periods is $\lambda_1(h, 2) = \{R_1, R_2\}$, whereas $\lambda_2(h, 2) = \{C_2\}$. As to orders, player 1's action R_1 is encountered first, then R_2, so $o_{1,h}(1) = R_1, o_{1,h}(2) = R_2$. Similarly, $o_{2,h}(1) = C_2, o_{2,h}(2) = C_1$. $\triangleleft$

3. Adaptive play

This section presents a class of Markov chains to model 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 $T \in \mathbb{N}$.[5] Consequently, we define the ***state space*** H to consist of all histories $h = (a^1, \ldots, a^L)$ with length at least T, i.e., $h \in \cup_{K \in \mathbb{N}, K \geq T} A^K$.

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 $h \in H$. To do so, we model beliefs and responses to them.

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

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

[5] 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. For the convergence result, the exact probabilities are irrelevant: what matters is that some are positive, others zero. We therefore refrain from restricting attention to specific belief formation processes or updating procedures. As long as beliefs are sufficiently diverse — see Remark 3.2 or the related discussion in Hurkens (1995, pp. 310–311) — it is immaterial how they are formed.

RESPONSES: Given a belief $\alpha_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(\lambda_j(h,T))$, we assume that player i chooses the most recent best reply to α_{-i} if such a best reply exists, i.e., if in state h some best reply to α_{-i} has been played before. Otherwise, player i chooses each best reply to α_{-i} with positive probability, i.e., it is drawn from a probability measure $\mathbb{P}_{\alpha_{-i}}$ over A_i whose support coincides with the set of best replies $BR_i(\alpha_{-i})$. Players thus have a bias towards recent choices.[6]

Together, the probability distributions $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$ that fix 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 $(h,h') \in H \times H$. If $P(h,h') > 0$, then histories $h, h' \in H$ satisfy conditions P1 and P2 in Fig. 2.

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, firstly, 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 $\times_{j \in N \setminus \{i\}} \Delta(\lambda_j(h,T))$. Secondly, it models the bias towards recent choices: whenever possible, each player $i \in N$ chooses the most recent best reply to belief α_{-i}.

Let $\mathscr{P}$ be the class of transition probability functions P achieved in this way, i.e., using probability distributions $\{\mathbb{P}_{(i,(a^1,\ldots,a^T))} \mid i \in N, (a^1, \ldots, a^T) \in A^T\}$ and the behavioral bias, and with $P(h,h') > 0$ if and only if states $h, h' \in H$ satisfy conditions P1 and P2 in Fig. 2.

Remark 3.1 The behavioral bias towards recent choices modeled in P2 requires that a player with multiple best replies against his current belief recalls whether one of them was

[6]The probability of choosing $a_i \in A_i$ against beliefs to which no best reply was chosen before is

$$\int_{\{\alpha_{-i} \mid BR_i(\alpha_{-i}) \cap \lambda_i(h,L) = \emptyset\}} \mathbb{P}_{\alpha_{-i}}(a_i)\, \mathrm{d}\mathbb{P}_{(i,(a^1,\ldots,a^T))},$$

i.e., $\alpha_{-i} \mapsto \mathbb{P}_{\alpha_{-i}}(a_i)$ is assumed to be Borel measurable.

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

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

played more recently than another. However, this does not require players to have perfect memory about their own actions: If you played one best reply yesterday and another a week ago, your choice is independent of whether you also adopted these actions further away in the past. All that matters is that each player $i \in N$ in history $h \in H$ recalls the order $o_{i,h}$ defined in Section 2. This is a considerably more modest requirement than remembering the entire history of own actions: $o_{i,h}$ specifies a simple linear order of at most $|A_i|$ actions. Between consecutive rounds of play, this linear order either remains the same or changes in the following way: The action ranked first (the most recent action) is changed and the other actions are moved one step down the ladder. For instance, even after numerous rounds of play, the only thing a player with just two actions needs to recall from his own past is last period's action. $\triangleleft$

Remark 3.2 Inherent in the definition of the class $\mathscr{P}$ of transition probability functions (see Fig. 2) is that beliefs must be "sufficiently diverse" to assure that player $i \in N$ has a positive probability of selecting $a_i \in A_i$ whenever it is a (most recent) best reply to some belief over recent past play. More specifically, by P2, player i is tempted to play a_i against beliefs α_{-i} over recent past play to which it is the most recent best reply or — if no such most recent best reply exists — to which it is an arbitrary best reply. If the set of such "tempting" beliefs is nonempty, player i assigns positive probability to it. $\triangleleft$

For each $k \in \mathbb{N}$, $P^k : H \times H \to [0, 1]$ denotes the k-step transition probabilities of the Markov process with transition probability function $P \in \mathscr{P}$: $P^1 = P$ and $P^k = P \circ P^{k-1}$ if $k > 1$.

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 beliefs are 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 Appendix A.

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 of G if the probability that the process after k steps is in a state $h \in H$ where

- the most recently played action profile $\ell(h)$ lies in some minimal prep set X of G:

$$\ell(h) \in X$$

- all future action profiles remain inside X:

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

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*

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

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

Remarks:
(i): If the game has several minimal prep sets, the one selected by the learning process typically depends on initial conditions. For instance, if the initial state is such that the collection of most recent actions is a minimal prep set X, the process settles down in X.
(ii): Condition P2 assures that play will not settle down in proper subsets of a minimal prep set. Suppose play settles down in a product set Y properly contained in a minimal prep set X. Since Y is not a prep set, there is a player i with a belief over recent past play against which Y_i contains no best reply. Condition P2 assures that player i with positive probability chooses such a best reply, i.e., an action outside Y_i, contradicting the

assumption that play has settled down in Y. A similar intuition is used in the proof of the Theorem in Appendix A (Lemma A.1).
(iii): The statement of Theorem 4.1 follows the traditional pattern (cf. Hurkens, 1995, Fudenberg and Levine, 1998, and Young, 1998): if memory is 'sufficiently long', play settles down in sets of a certain type. Thus, we indicate a sufficient length in (1), without aiming at sharpness. Were one to exploit specific features of a game, the bound might be decreased (e.g. Kets and Voorneveld, 2007).

Steps of the proof: 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 $\times_{i \in N} \lambda_i(h_1, T) \subseteq A$ of actions played in the past T periods is a prep set.

The intuition behind this step is as follows. If, for some state $g \in H$ and some $k \leq T$, the product set $\times_{i \in N} \lambda_i(g, k)$ is a prep set, then with positive probability, players choose actions from this prep set for $T-k$ periods in a row. If on the other hand, $\times_{i \in N} \lambda_i(g, k)$ is not a prep set, then there is a nonempty set of players $i \in N$ with a belief $\alpha^*_{-i} \in \times_{j \in N \setminus \{i\}} \Delta(\lambda_j(g, k))$ over play in the past k periods to which $\lambda_i(g, k)$ 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 $\times_{i \in N} \lambda_i(g_k, k)$ is strictly increasing with respect to set inclusion (see Lemma A.1 in Appendix A). All these sets are contained in the finite set A of action profiles which is a 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 $\times_{i \in N} \lambda_i(g_K, K)$ 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 $X := \times_{i \in N} \lambda_i(h_2, T)$ is a minimal prep set.

Indeed, let $X = \times_{i \in N} X_i \subseteq \times_{i \in N} \lambda_i(h_1, T)$ 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 $i \in N$, $\lambda_i(h', |X_i|) = X_i$, 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 within a finite number of steps such that $\times_{i\in N}\lambda_i(h_2,T) = X$ 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 state h_2, $\times_{i\in N}\lambda_i(h_2,T) = X$ 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 a player i may have about 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, Step 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 finitely many steps. Since the initial history was chosen arbitrarily, this eventually happens with probability one, finishing the proof.

5. 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 A) 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 $\times_{i\in N}\lambda_i(h_2,T)$ 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, whenever possible, choose the most recent best reply to a certain belief. However, any class of adjustment processes that respects these conditions on the sign of the transition probabilities will converge to minimal prep sets. Hence, one can easily extend the class of adjustment processes with this limit behavior.

In particular, suppose that for each player $i \in N$, the response to a belief drawn from recent past play is chosen 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' used actions appear in h. That is, for each pair of states

$h = (a^1, \ldots, a^L), g = (b^1, \ldots, b^K) \in H$:

$$\left.\begin{array}{rcl}(a^1, \ldots, a^T) & = & (b^1, \ldots, b^T) \\ o_{i,h} = o_{i,g} & & \text{for all } i \in N\end{array}\right\} \Rightarrow R_{i,h} = R_{i,g} \text{ for all } i \in N. \tag{2}$$

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. 2) 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 $\widetilde{\mathscr{P}}$ be the collection of transition probability functions $\{P_R : H \times H \to [0, 1] \mid R = (R_{i,h})_{i \in N, h \in H}\}$ satisfying the restrictions on the sign of the transition probabilities instrumental to the proof of Theorem 4.1, i.e., 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 $X := \times_{i \in N} \lambda_i(h, k)$ is a minimal prep set for some $k \geq T$ and $P_R(h, h') > 0$, then $\times_{i \in N} \lambda_i(h', k+1) = X$, i.e., $\ell(h') \in X$.

One easily shows that $\mathscr{P} \subseteq \widetilde{\mathscr{P}}$. In many games, the set inclusion is strict and one finds processes in $\widetilde{\mathscr{P}} \setminus \mathscr{P}$ by letting players choose more freely among recent best replies. By construction, if memory is sufficiently long, processes in $\widetilde{\mathscr{P}}$ eventually settle down in minimal prep sets. To summarize (the proof is in Appendix B):

Proposition 5.1 *Let $G = \langle N, (A_i)_{i \in N}, (u_i)_{i \in N} \rangle$ be a game and let $T \in \mathbb{N}$. 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.*

Finally, we show why it is essential for convergence to minimal prep sets that players keep track of the order $o_{i,h}$ defined in Section 2, rather than just the order of their actions over the past T rounds of play.

Remark 5.2 If players remember from their own past only their actions in the previous T periods, the resulting processes need not converge to minimal prep sets: None of the players $i \in N$ can condition his behavior on the order $o_{i,h}$ of actions chosen more than T

periods ago at state h. To see why this prevents convergence to minimal prep sets, refer back to Fig. 1. Suppose that over the past T rounds, players have chosen the actions from minimal prep set $X = \{R_1, R_2\} \times \{C_1, C_2\}$. Why wouldn't play settle down there? Suppose players play (R_1, C_2) at a given round, which are best replies to beliefs over X. In response to these actions, there is a positive probability that they choose (R_2, C_2) for T consecutive periods. At that point, player 2's only feasible belief over past play is that player 1 chooses R_2. Player 2 recalls only his past T actions, i.e., just C_2 which is no best reply to R_2. Therefore, he chooses among the best replies $\{C_1, C_3\}$ to R_2, which means that he may jump outside the minimal prep set X by selecting C_3. $\triangleleft$

6. Concluding remarks

The purpose of this paper was to study discrete-time best-response processes with a behaviorally plausible bias towards recent actions. Such processes were shown to settle down in minimal prep sets. This dynamic motivation complements earlier papers on minimal prep sets in a static environment, where the concept is compared with many other solution concepts (Voorneveld, 2004, 2005, Voorneveld et al., 2005) and shown to have genuine "bite" in economic applications (Tercieux and Voorneveld, 2005), even in cases where, for instance, minimal curb sets have no cutting power whatsoever.

Several modifications of these processes were discussed in the previous section. We cannot possibly do justice to the long list of choice biases discussed in the behavioral economics literature. An interesting direction for future research — although outside the scope of the current paper — would be to more systematically investigate the links between different types of behaviorally plausible biases in adjustment processes and the corresponding limiting behavior.

Acknowledgments

Financial support from the Netherlands Organization for Scientific Research (NWO) and the Wallander/Hedelius Foundation is gratefully acknowledged. We thank Kaushik Basu, Wieland Müller, Hans Peters, Dolf Talman, Dries Vermeulen, Jörgen Weibull, and several audiences for helpful comments and discussions. We thank two referees and an associate editor of this journal for their careful reading and very concrete suggestions for improvements.

Appendix A 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 $T \geq \max\{\sum_{i\in N} |A_i| - n + 1, 2|A_1|, \ldots, 2|A_n|\}$, 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 $a_i \in \lambda_i(h, T)$ is ***blocked in*** h if there is no belief $\alpha_{-i} \in \times_{j\in N\setminus\{i\}}\Delta(\lambda_j(h, T))$ against which it is the most recent best reply. Finally, the ***frequency*** with which player i chose action $a_i \in \lambda_i(h, T)$ during the past T rounds of history h is

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

We now prove the four steps of Theorem 4.1.

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 $\times_{i\in N}\lambda_i(h_1, T) \subseteq A$ 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 $\times_{i\in N}\lambda_i(h, t) \subseteq A$ is not a prep set. Then the process moves with positive probability to a successor h' of h where

$$\times_{i\in N}\lambda_i(h, t) \subset \times_{i\in N}\lambda_i(h', t + 1). \tag{3}$$

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

$$\times_{i\in N}\lambda_i(h, t) = \times_{i\in N}\lambda_i(h', t + 1). \tag{4}$$

Proof. **(a):** Since $\times_{i\in N}\lambda_i(h,t) \subseteq A$ is not a prep set, there is a nonempty set $S \subseteq N$ of players $i \in N$ with a belief $\alpha^*_{-i} \in \times_{j\in N\setminus\{i\}} \Delta(\lambda_j(h,t))$ over the play in the past t periods to which $\lambda_i(h,t)$ does not contain a best reply: $BR_i(\alpha^*_{-i}) \cap \lambda_i(h,t) = \emptyset$. Fix such a belief α^*_{-i} for each $i \in S$ and let $b_i \in BR_i(\alpha^*_{-i})$ be a best reply to α^*_{-i} chosen in accordance with P2: it is the most recent one if $BR_i(\alpha^*_{-i}) \cap \{a_i^1, \dots, a_i^L\} \neq \emptyset$. For each $i \in N \setminus S$, let $b_i \in \lambda_i(h,t)$ 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 $h' = (b, a^1, \dots, a^L)$. Now (3) holds by construction: if $i \in N \setminus S$, then $b_i \in \lambda_i(h,t)$, so $\lambda_i(h,t) = \lambda_i(h', t+1)$, and if $i \in S$, then $b_i \notin \lambda_i(h,t)$, so $\lambda_i(h,t) \subset \lambda_i(h,t) \cup \{b_i\} = \lambda_i(h', t+1)$.
(b): Fix, for each $i \in N$, a belief $\alpha_{-i} \in \times_{j\in N\setminus\{i\}} \Delta(\lambda_j(h,t))$ over the play in the past t periods. Since $\times_{i\in N}\lambda_i(h,t)$ is a prep set, there is an action $b_i \in \lambda_i(h,t)$ which is the most recent best reply to this belief. By P1 and P2, the process moves with positive probability from h to $h' = (b, a^1, \dots, a^L)$. Since $b_i \in \lambda_i(h,t)$ for all $i \in N$, it follows that $\lambda_i(h', t+1) = \lambda_i(h,t)$, so (4) holds. □

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

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

with strict inclusion if $\times_{i\in N}\lambda_i(g_k, k)$ is not a prep set and equality otherwise. The sequence of product sets $\times_{i\in N}\lambda_i(g_k, k)$ 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; $\times_{i\in N}\lambda_i(g_1, 1)$ 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 $\times_{i\in N}\lambda_i(g_{K+1}, K+1)$ is a prep set.[7] In the final $T-K-1$ steps, we proceed to a state g_T, where

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

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

[7] This motivates the term $M := \sum_{i\in N} |A_i| - n + 1$ in the lower bound on T in (1): reaching a prep set can take $M-1$ steps; recalling the added actions and those in g_1 can consequently take a memory length M.

A.2 States without blocked actions

In this section, we show that from a state $h \in H$ such that $\times_{i\in N}\lambda_i(h,T)$ is a prep set, the process moves with positive probability within a finite number of steps to a state $h' \in H$ where $\times_{i\in N}\lambda_i(h',T) \subseteq \times_{i\in N}\lambda_i(h,T)$ is a prep set without blocked actions. This is established in Lemma A.3, using Lemma A.2. Furthermore, in Lemma A.4 we show that when considering a sequence $g_1,\ldots,g_K$ such that, for all $k = 1,\ldots,K$, $\times_{i\in N}\lambda_i(g_k,T)$ is a prep set and $\times_{i\in N}\lambda_i(g_1,T) \supseteq \cdots \supseteq \times_{i\in N}\lambda_i(g_K,T)$, 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 of the following subsections.

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

- *if $\tau_i(h)$ is blocked, let $\beta_i(h) \in \lambda_i(h,T)$ 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:

$$P(h,h') > 0 \tag{5}$$
$$\times_{i\in N}\lambda_i(h',T) \subseteq \times_{i\in N}\lambda_i(h,T) \tag{6}$$
$$\times_{i\in N}\lambda_i(h',T) \quad \textit{is a prep set.} \tag{7}$$

Proof. For all $i \in N$, $\beta_i(h) \in \lambda_i(h,T)$ is not blocked by definition: there is a belief $\alpha_{-i} \in \times_{j\in N\setminus\{i\}}\Delta(\lambda_j(h,T))$ against which $\beta_i(h)$ is the most recent best reply. By P1 and P2, (5) holds. Since $\beta_i(h) \in \lambda_i(h,T)$ for all $i \in N$, (6) holds. To prove (7), let $i \in N$ and $\alpha_{-i} \in \times_{j\in N\setminus\{i\}}\Delta(\lambda_j(h',T))$. To show: $BR_i(\alpha_{-i}) \cap \lambda_i(h',T) \neq \emptyset$. By construction, $\lambda_i(h',T)$ equals either $\lambda_i(h,T)$ or, if $\tau_i(h)$ was blocked and chosen only once in the most recent T periods of history h, $\lambda_i(h,T) \setminus \{\tau_i(h)\}$. Consequently, $\lambda_i(h',T)$ still contains a best reply to every belief over $\times_{j\in N\setminus\{i\}}\Delta(\lambda_j(h,T))$, in particular to every belief over the subset $\times_{j\in N\setminus\{i\}}\Delta(\lambda_j(h',T))$. □

Claim (6) 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 $a_i \in \lambda_i(h',T)$, then it remains blocked:

$$\text{if } a_i := \tau_i(h) \text{ was blocked in } h \text{ and } a_i \in \lambda_i(h',T), \text{ then it is blocked in } h'. \tag{8}$$

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 $\times_{i \in N} \lambda_i(h', T)$, in which case (6) holds with strict inclusion, or it remains blocked in h' by (8), 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 $g_{k+1} = (\beta(g_k); g_k)$ for all $k \in \mathbb{N}$, 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

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

By (7), 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 $\times_{i \in N} \lambda_i(h, T)$ 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 $\times_{i \in N} \lambda_i(h', T) \subset \times_{i \in N} \lambda_i(h, T)$ is a prep set and h' contains no blocked actions.*

The proof of Step 2 uses so-called drag-to-front operations (Section A.3) and neighbor switches (Section A.4) to establish the following: Given a state $g_1 \in H$ where $\times_{i \in N} \lambda_i(g_1, T)$ is a prep set, the process moves with positive probability in a finite number of steps through a sequence of states $g_1, g_2, \ldots, g_K$ such that

$$\forall k = 1, \ldots, K : \times_{i \in N} \lambda_i(g_k, T) \text{ is a prep set,} \tag{9}$$

$$\times_{i \in N} \lambda_i(g_1, T) \supseteq \times_{i \in N} \lambda_i(g_2, T) \supseteq \cdots \supseteq \times_{i \in N} \lambda_i(g_K, T), \tag{10}$$

and g_K has the property that for some minimal prep set $X = \times_{i \in N} X_i$ and each $i \in N$:

$$\lambda_i(g_K, |X_i|) = X_i,$$

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. If any of the states g_k contains a blocked action, apply Lemma A.3 to move to a state g' where $\times_{i \in N} \lambda_i(g', T) \subset \times_{i \in N} \lambda_i(g_k, T)$ is a prep set and g' contains no blocked actions. Then, we can start the repeated use of drag-to-front operations and neighbor switches anew from g'. Since there are only finitely many prep sets and the prep set $\times_{i \in N} \lambda_i(g', T)$ is strictly contained in $\times_{i \in N} \lambda_i(g_k, T)$, we eventually reach in a finite number of steps a state from which we can apply drag-to-front operations and neighbor switches without ever encountering a state with a blocked action. Hence:

Lemma A.4 *In a sequence of states $(g_k)_{k=1,\dots,K}$ satisfying (9) and (10), obtained using drag-to-front operations and neighbor switches, we may assume w.l.o.g. that none of the states contains a blocked action.*

A.3 Drag-to-front operations

Consider a state $h \in H$ containing no blocked actions for which $\times_{i\in N}\lambda_i(h,T)$ is a prep set. Then, by definition, for each $i \in N$, $\beta_i(h) = \tau_i(h)$, the action player i chose T periods ago in state h (see Lemma A.2). 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 $a_j \in \lambda_j(h,T)$ with frequency $f_j(h,a_j) = 1$. Since[8] $T \geq 2|A_j|$ by (1), there must be an action $b_j \in \lambda_j(h,T)$ with frequency $f_j(h,b_j) \geq 3$. By Lemma 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 $\gamma_i(h) \in \lambda_i(h,T)$ as follows:

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

Set $h' = (\gamma(h);h)$ with $\gamma(h) = (\gamma_i(h))_{i\in N}$. Recall: (1) $\gamma_i(h) \in \lambda_i(h,T)$ for all $i \in N$, (2) $\times_{i\in N}\lambda_i(h,T)$ 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 $\alpha_{-i} \in \times_{k\in N\setminus\{i\}}\Delta(\lambda_k(h,T))$. By P1 and P2, $P(h,h') > 0$.

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

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

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.

[8] This motivates the term $2|A_j|$ in the lower bound on T. If the memory length is below this bound, neighbor switches as defined in Section A.4 may cause actions from prep sets to disappear from a player's recollection.

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

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

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

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

A.4 Neighbor switches

Repeatedly applying drag-to-front operations starting in a state $h \in H$ where no actions are blocked and $\times_{i \in N} \lambda_i(h, T)$ 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.

Instead, it is possible that some player i chooses his actions in such a way that the process moves to a state in which the order in which player i plays two neighboring actions — say those chosen t and $t + 1$ periods ago in state h — is changed, while the others continue to play actions in their given order. For instance, the process may move from Fig. 3.a to Fig. 3.e, where player i's order of actions b and c, chosen 2 and 3 periods ago in Fig 3.a, respectively, is reversed while the order of actions of players $j \neq i$ is unchanged. 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:[9] use drag-to-front operations until the actions to be switched 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 that are to be switched 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

[9]Fig. 3 is for illustration only; we assume that all steps we describe there are feasible.

instead of b). Finally, use drag-to-front operations until the switched actions are again at coordinates t and $t+1$ in the recent past play (the transition from Fig. 3.d to Fig. 3.e). Formally:

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 b and c, keeping those of players $j \neq i$ in the same order.

Lemma A.6 *Let $h \in H$ satisfy* [C1] *to* [C3]*. Let $i \in N, t \in \{1, \ldots, T-1\}$. Assuming w.l.o.g. (Lemma A.4) that we encounter no blocked actions, the process moves with 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 $j = i$ and $k \notin \{t, t+1\}$, or if $j \neq i$, 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 g_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 g_2 of g_1 as follows: for each $j \in N \setminus \{i\}$, set $s_j^1 = \tau_j(g_1)$ and set $s_i^1 = b_i$. Define $g_2 = (s^1; g_1)$, where $s^1 = (s_j^1)_{j \in N}$.

Construct a successor g_3 of g_2 as follows: for each $j \in N \setminus \{i\}$, set $s_j^2 = \tau_j(g_2)$ and set $s_i^2 = a_i$. Define $g_3 = (s^2; g_2)$, 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 g_1 to g_2, i chooses b_i, in going from g_2 to g_3, i chooses a_i, rather than playing first a_i, then b_i.

As $\times_{i \in N} \lambda_i(g_1, T)$ is a prep set and no actions are blocked in g_1, it follows from P1 and P2 that $P(g_1, g_2) > 0$. Moreover, as all actions in h have frequency at least 2, we have that $\lambda_i(g_1, T) = \lambda_i(g_2, T)$ for all $i \in N$. Hence, also $\times_{i \in N} \lambda_i(g_2, T)$ is a prep set. By Lemma A.4 we may assume that g_2 contains no blocked actions. Hence, also $P(g_2, g_3) > 0$. Moreover, it is easy to see that frequencies in g_3 are identical to frequencies in g_1, i.e., at least equal to 2. We can thus conclude that also g_3 satisfies [C1] to [C3].

In g_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 Steps 2 to 4

Step 2: Let $h_1 \in H$ be such that $\times_{i \in N} \lambda_i(h_1, T)$ is a prep set. The process moves with positive probability in a finite number of steps to a state $h_2 \in H$ where $\times_{i \in N} \lambda_i(h_2, T)$ 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 $X = \times_{i \in N} X_i \subseteq \times_{i \in N} \lambda_i(g, T)$ be a minimal prep set. Assuming w.l.o.g. (Lemma 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 $i \in N$, $\lambda_i(g_0, |X_i|) = X_i$, 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 $g_k := ((a_i(g_{k-1}, |X_i|))_{i \in N}; g_{k-1}) \in H$, 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

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

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

Step 3: Let $h_2 \in H$ be such that $X = \times_{i \in N} \lambda_i(h_2, T)$ is a minimal prep set. After reaching

h_2, all action profiles that are played with positive probability lie in X:

$$\forall k \in \mathbb{N}, \forall h \in H : P^k(h_2, h) > 0 \Rightarrow \ell(h) \in X. \tag{11}$$

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 $\times_{i \in N} \lambda_i(h_2, T)$ 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 $h \in H$:

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

and hence

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

In particular, this means $\ell(h) \in X$, i.e., (11) 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. Call two states $h = (a^1, \ldots, a^L)$ and $g = (b^1, \ldots, b^K)$ in H equivalent, denoted $h \sim g$, if they have the same account of recent past play and the same order in which each player i's actions are encountered:

$$h \sim g \Leftrightarrow \begin{cases} (a^1, \ldots, a^T) = (b^1, \ldots, b^T), \\ o_{i,h} = o_{i,g} \text{ for all } i \in N. \end{cases}$$

Notice that $\sim$ is an equivalence relation on H; for each $h \in H$, let $[h] = \{h' \in H : h \sim h'\}$ be the equivalence class containing h. Recall from Section 3 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, firstly, player i's beliefs α_{-i} are drawn from a probability distribution $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$ and, secondly, his response is (whenever possible) the most recent best reply to this belief or (otherwise) drawn from a probability distribution $\mathbb{P}_{\alpha_{-i}}$ over his best replies. Thus, player i's choice behavior is the same in two equivalent states. Since there are only finitely many elements in A^T and N, it follows that the set of positive transition probabilities $\{P(h, h') \mid h, h' \in H, P(h, h') > 0\}$ 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. □

Appendix B Proof of Proposition 5.1

Proof. [**The inclusion** $\mathscr{P} \subseteq \widetilde{\mathscr{P}}$**:**] Let $P \in \mathscr{P}$. The probability $R_{i,h}(a_i)$ that player $i \in N$ in state $h = (a^1, \ldots, a^L) \in H$ chooses action $a_i \in A_i$ equals the probability of drawing a belief α_{-i} from $\mathbb{P}_{(i,(a^1,\ldots,a^T))}$ to which:

(i) a_i is the most recent best reply, or, alternatively,

(ii) no best reply was played before, but response a_i is drawn from $\mathbb{P}_{\alpha_{-i}}$.

Hence, 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 $P \in \widetilde{\mathscr{P}}$.

[**Convergence:**] The proof of Theorem 4.1 in Appendix A 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 (2), 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. □

References

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

Crawford, V. P., 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.

Kets, W., Voorneveld, M., 2007. Congestion, equilibrium and learning: The minority game. CentER Discussion Paper 2007-61.

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 Conflict. Cambridge MA: Harvard University Press.

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.

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.

Voorneveld, M., Kets, W., Norde, H., 2005. An axiomatization of minimal curb sets. Int. J. Game Theory 33, 479-490.

Young, H.P., 1998. Individual Strategy and Social Structure. Princeton NJ: Princeton University Press.