JOURNAL OF ECONOMIC THEORY **82**, 65–88 (1998)
ARTICLE NO. ET982419

A Learning Approach to Auctions*

Shlomit Hon-Snir† and Dov Monderer‡

Faculty of Industrial Engineering and Management,
Technion–Israel Institute of Technology, Haifa 32000, Israel

and

Aner Sela§

Mannheim University, Sonderforschungsbereich 504, L 13, 15, 68131 Mannheim, Germany

Received November 17, 1996; revised February 23, 1998

We analyze a repeated first-price auction in which the types of the players are determined before the first round. It is proved that if every player is using either a belief-based learning scheme with bounded recall or a generalized fictitious play learning scheme, then after sufficiently long time, the players' bids are in equilibrium in the one-shot auction in which the types are commonly known. *Journal of Economic Literature Classification Numbers*: C72, C73, D44, D83.

Contents

* First version: september 1996. An Internet version of this paper can be found in the Economics Working Paper Archive http:// econwpa.wustl.edu (No. 9610004). We thank the participants of the microeconomics workshop, held at Northwestern University in August 1996 for their useful comments, and in particular we thank Larry Samuelson for Example 3. We also thank Ezra Einy, Doron Sonsino, Benyamin Shitovitz, anonymous associate editor, anonymous referee and especially Ron Holzman and Drew Fudenberg for helpful discussions. This work was supported by the Fund for the Promotion of Research in the Technion. Financial support from the Deutche Forschungsgemeinschaft, sonderforchungsbereich 303 at the university of Bonn is gratefully acknowledged.

† hon@tx.technion.ac.il.
‡ dov@ie.technion.ac.il.
§ aner@sfb504.uni.mannheim.de.

0022-0531/98 $25.00

1. INTRODUCTION

This work deals with a first-price sealed-bid auction of a single item. Such type of auction, as well as many other types, have been extensively used as selling mechanism and have been the subject of an intensive theoretical research in economics and operations research.[1] Much of this research has focused on the equilibrium analysis of the corresponding one-shot Bayesian game.[2] Other research efforts have been devoted to auction design, based on equilibrium analysis.[3] This work differs from previous ones in three main aspects: (a) it discusses discrete models, (b) it deals with repeated auctions with incomplete information,[4] and in particular, (c) it does not analyze the repeated-game equilibria set, but rather employs "learning theory." More precisely, we analyze the path generated by players who use various classes of belief-based learning schemes, including the class of learning schemes with bounded recall and the class of generalized fictitious play learning schemes. Roughly speaking, a player with a recall of size m assigns a positive probability to a vector of the other players' bids if and only if this vector was used in one of the last m stages. A player that uses a generalized fictitious play learning scheme assumes that his opponents' next bid vector is distributed according to a weighted empirical distribution of their past bid vectors.[5] We further assume that the players are risk neutral and that each player's type is determined before the first auction and does not vary with time.[6] In our main result (Theorem A in

[1] It is very difficult to list the numerous papers on auctions. The reader is referred to to the surveys of Stark and Rothkopf [28], Milgrom [15, 16], McAfee and McMillan [12], Wilson [32], Wolfstetter [33], Laffont [9], and to the evolving recent literature concerning auctions on spectrum rights (e.g., McMillan [14] and Cramton [3]).

[2] See, e.g. (in addition to the above mentioned surveys), the early work of Vickrey [30] and the more recent and comprehensive approach to equilibrium of Milgrom and Weber [17].

[3] See Harris and Raviv [6], Myerson [22], and Riley and Samuelson [24].

[4] Various types of government bonds are repeatedly sold in first-price auctions. For other examples see, e.g., Ashenfelter [1].

[5] Note that the players in our model are not sophisticated. They do not attempt to learn their opponents' types or to hide their own types. They merely make a simple statistical inference about their opponents' next move. That is, our model does not exhibit the ratchet effect appearing in some of the equilibrium strategies in repeated auctions. Equilibrium analysis of repeated first-price auctions in the framework of repeated games with incomplete information is complex. Therefore this theory is restricted and not conclusive. The reader is referred to Laffont [9] for a survey of the relevant literature. An analysis of repeated second-price auctionswith incomplete information is given in Bikhchandani [2]. Sequential auctions in which each player wishes to purchase at most one unit are well-understood and discussed, e.g., in Milgrom and Weber [17], Weber [31] and McAfee and Vincent [13]. Finally, McAfee [11] discusses a dynamic setup for general auction mechanisms. He uses a solution concept that involves elements of competitive equilibrium and strategic equilibrium.

[6] See Section 7 for a discussion of this assumption.

Section 5) we prove under mild assumptions concerning tie-breaking rules that after sufficiently long time the players play an equilibrium of the one-shot auction in which players' types are common knowledge. This result means, that generically, the player with the highest valuation wins the object and pays the second-highest valuation. That is, under our belief-based learning assumption, a repeated first price auction yields in the long-run the outcome of a one-shot second-price auction. In Section 6 we show by examples that Theorem A does not hold when we remove one of the tie-breaking rules. However, we show that for two-person auctions, even without this tie-breaking rule, if both players use a generalized fictitious play learning scheme, then the beliefs of the players approach a mixed-action equilibrium.This last result does not hold when the players use a learning scheme with bounded recall as is shown by an example. Section 7 is devoted to some other learning schemes that seem natural in the context of auctions, and in both Section 6 and Section 7 we provide some remarks and open problems. The proofs of the main theorems are given in Section 8.

2. REPEATED (DISCRETE) FIRST-PRICE SEALED-BID AUCTION

Let $N=\{1, 2, ..., n\}$ be the set of players. In the one-shot auction $A(v^1, v^2, ..., v^n)$, Player i has a type v^i which is a positive integer. That is, v^i is the expected monetary value of the item for Player i. The action set of each player is the set $Z_+=\{1, 2, ...\}$ of positive integers. When every player $i \in N$ makes a bid $x^i \in Z_+$, the player with the maximal bid wins the object. If there is more than one such player, we deviate from the standard theory by purifying the game: Instead of assuming that in such a case the winner is determined by a lottery, we assume that every winner receives his expected utility. That is, if $x^i = x^{\max} = \max_{j \in N} x^j$, then i receives $(1/w)(v^i - x^i)$, where w denotes the number of players j with $x^j = x^{\max}$. Our purifying method is harmless if the players are assumed risk neutral, as we indeed assume. The types are selected by Nature according to the probability distribution λ over $(Z_+)^N$. Every player knows his type. The precise nature of λ as well as the information channels of the players are not important for the "learning" analysis (amthough they play a crucial role in the standard equilibrium analysis). In the repeated auction, Nature chooses the types $v^1, v^2, ..., v^n$, and the auction is repeatedly conducted. The repeated auction is denoted by $RA(v^1, v^2, ..., v^n)$.

3. BELIEF-BASED LEARNING

Consider a repeated game in strategic form. The one-shot game is denoted by G. The set of players in G is $N=\{1, 2, ..., n\}$. The action set of

Player i is S^i, and i's utility function is $u^i: S \to R$, where $S = \times_{j \in N} S^j$ and R denotes the set of real numbers.[7] Let $H_t = S^t$ be the set of histories of length t. By convention, H_0 is a singleton. A *strategy* for player i is a function $f^i: \bigcup_{t=0}^{\infty} H_t \to S^i$. For a finite set X, $\Delta(X)$ denotes the set of probability measures over X. A *belief function* for i is a function $B^i: \bigcup_{t=e^i}^{\infty} H_t \to \Delta(S^{-i})$, where e^i is a given positive integer, and $S^{-i} = \times_{j \neq i} S^j$. $B^i(h_{t-1})$ is the belief of Player i about the tth joint action of all other players, after he observes the history $h_{t-1} = (x_1, ..., x_{t-1})$. Player i generates beliefs only after he observes at least e^i action profiles.[8] Let BR^i be the pure best response correspondence of Player i at the one-shot game. A *learning scheme* for i is a pair (B^i, f^i) such that $f^i(h_{t-1}) \in BR^i(B^i(h_{t-1}))$ for every $t > e^i$. We deal with learning schemes that satisfy stronger conditions than those of Milgrom and Roberts [18]—for an infinite history $h = (x_1, x_2, ...)$, we denote $(x_1, ..., x_t)$ by $h_{[t]}$.For two finite histories $h = (x_1, x_2, ..., x_t) \in H_t$ and $h' = (z_1, z_2, ..., z_{t'}) \in H_{t'}$ we denote by (h, h') the history $(x_1, x_2, ..., x_t, z_1, z_2, ... z_{t'})$ in $H_{t+t'}$. A learning scheme (B^i, f^i) is *adaptive* if it satisfies the following three conditions for every infinite history $h = (x_1, x_2, ...)$:

AD1. For every $\varepsilon > 0$ and for every $T > e^i$ there exists a positive integer M, such that for every $s > M$ for every $x^{-i} \in S^{-i}$ and for every $h_s = (z_1, ..., z_s) \in H_s$ if $z_t^{-i} \neq x^{-i}$ for every $1 \leqslant t \leqslant s$, then

$$B^i(h_{[T-1]}, h_s)(x^{-i}) < \varepsilon.$$

AD2. For $t > e^i$ if $B^i(h_{[t-1]})(x^{-i}) > 0$, then there exists $1 \leqslant s \leqslant t-1$ such that $x_s^{-i} = x^{-i}$.

AD3. For $x^{-i} \in S^{-i}$ and for $t > e^i$ if $x_{t-1}^{-i} = x^{-i}$, then $B^i(h_{[t-1]}) \times (x^{-i}) > 0$.

Condition AD1 means that Player i assigns a low probability to action profiles that have not been used for a long time. AD2 means that Player i assigns a 0-probability to a profile of actions that has never been used. Condition AD3 means that Player i does not ignore recent information.

[7] Actually we assume that G has been already chosen by Nature according to some probability measure λ. Each Player has received his signal and therefore knows the set of players, the set of actions $(S^j)_{j \in N}$ and its own utility function. The exact nature of λ is not relevant, however it is implicitly assumed that i's utility function which may depend on his signal, does not depend on the other players' signals. That is, the players can compute their best-response correspondences.

[8] e^i can be interpreted as the length of the experimenting period. The collection $(e^j)_{j \in N}$ can be used to capture the concept of prior beliefs.

(B^i, f^i) is a *fictitious play* (*FP*) *learning scheme* if for every $t > e^i$ and for every history h_{t-1}:

$$B^i(h_{t-1})(x^{-i}) = \frac{1}{t-1} \# \{1 \leqslant s \leqslant t-1: x_s^{-i} = x^{-i}\} \qquad \text{for every} \quad x^{-i} \in S^{-i}.$$

For a sequence $(w_t)_{t=1}^{\infty}$ of real numbers and for a subset A of positive integers we denote $w(A) = \sum_{t \in A} w_t$; $w(\{1, 2, ..., T\})$ is denoted by $w(T)$. (B^i, f^i) is a *generalized fictitious play* (*GFP*) *learning scheme* if there exists a nondecreasing sequence $w = (w_t)_{t=1}^{\infty}$ of positive real numberssuch that for every $t > e^i$ and for every history $h_{t-1} = (x_1, ..., x_{t-1})$:

$$B^i(h_{t-1})(x^{-i}) = \frac{1}{w(t-1)} \sum_{\{1 \leqslant s \leqslant t-1: x_s^{-i} = x^{-i}\}} w_s \qquad \text{for every} \quad x^{-i} \in S^{-i}.$$

If $w_t = 1$ for every $t \geqslant 1$, then the associated learning scheme is the FP learning scheme. Note that every GFP learning scheme is adaptive, because $\lim_{s \to \infty}(w(T)/w(T+s)) = 0$. We say that a learning scheme (B^i, f^i) has a bounded recall if there exists $1 \leqslant m^i \leqslant e^i$ such that the following two conditions are satisfied:

BR1. For $t > e^i$ both B^i and f^i depend only on the last m^i action profiles. That is, for every $(z_1, ..., z_{m^i}) \in H_{m^i}$ and for every h_{t-m^i-1}, $\bar{h}_{t-m^i-1} \in H_{t-m^i-1}$,

$$B^i(h_{t-m^i-1}, z_1, ..., z_{m^i}) = B^i(\bar{h}_{t-m^i-1}, z_1 ... z_{m^i})$$

and

$$f^i(h_{t-m^i-1}, z_1, ..., z_{m^i}) = f^i(\bar{h}_{t-m^i-1}, z_1 ... z_{m^i}).$$

BR2. For $t > e^i$, $B^i(h_{t-1})(x^{-i}) > 0$ if and only if x^{-i} was used in one of the last m^i stages (that is, if and only if there exists $t - m^i \leqslant s \leqslant t-1$ such that $x_s^{-i} = x^{-i}$, where $h_{t-1} = (x_1, ..., x_{t-1})$).

Note that Condition BR2 excludes degenerate learning schemes with zero recall, that is, learning schemes (B^i, f^i) for which there exists $p \in \Delta(S^{-i})$ such that for every $t > e^i$ and for every history h_{t-1}, $B^i(h_{t-1}) = p$.

LEMMA 3.1. *Every learning scheme with bounded recall is adaptive.*

Proof. Obviously, BR2 implies AD2 and AD3. It also implies a stronger version of AD1: For every $\varepsilon > 0$ and for every $T > e^i$, we can take $M = m^i$. ∎

We say that the recall size of a learning scheme with bounded recall is m^i, if $m^i \geqslant 1$ is the minimal positive integer which satisfies BR1 and BR2. The following simple lemma gives a useful principle. The converse of principle has already been proved in other versions by Fudenberg and Kreps [4], by Monderer and Sela [20] (who call it the "stability principle"), and by Fudenberg and Levine [5].

LEMMA 3.2. *Consider a repeated game as described above. Let* $h = (x_1, x_2, \ldots)$ *be a path that is generated when each player i uses either a learning scheme with bounded recall or a GFP learning scheme. Assume that there exists* $x \in S$ *and* T_0 *such that* $x_t = x$ *for every* $t \geqslant T_0$. *Then x is in equilibrium.*

Proof. If Player i uses a learning scheme with bounded recall, then for a certain large t he uses x^i as a best reply versus x^{-i}. If i uses a GFP learning scheme, then for every $\varepsilon > 0$, i uses x^i as a best reply to a belief which assigns to x^{-i} a probability greater than $1 - \varepsilon$. Therefore x^i must be a best reply to x^{-i}. ∎

Note that a belief function does not determine the learning scheme; If (B^i, f_1^i) and (B^i, f_2^i) are two learning schemes with the same belief function, then for $t > e^i$, $f_1^i(h_{t-1})$ may differ from $f_2^i(h_{t-1})$ for histories h_{t-1} for which $BR^i(B^i(h_{t-1}))$ is not a singleton. It is sometimes convenient to add a tie-breaking rule to the definition of a learning scheme. We will frequently use the following such rule:

TB1. If $t > e^i$ and $x_{t-1}^i \in BR^i(B^i(h_{t-1}))$, then $f^i(h_{t-1}) = x_{t-1}^i$.

Note that TB1 is only a partial tie-breaking rule. That is, there may be ties to which it is not applied.

4. BELIEF-BASED LEARNING IN AUCTIONS—MAIN THEOREM

We proceed to analyze the paths generated by players in the repeated auction $RA(v^1, v^2, \ldots, v^n)$, when each player uses a GFP learning scheme or a learning scheme with bounded recall. Note that for each Player i with $v^i > 1$, every bid $x^i \geqslant v^i$ is weakly dominated by the bid $v^i - 1$. We will assume the tie-breaking rule:

TB2. If $v^i > 1$, then Player i never chooses a bid exceeding $v^i - 1$.

THEOREM A. *Let* $RA(v^1, v^2, \ldots, v^n)$ *be a repeated first-price auction. Assume every player is using either a learning scheme with a bounded recall,* *or a GFP learning scheme, along with the tie-breaking rules TB1 and TB2. Then there exist a time T_0 and a strategy profile $x \in S$ which is in equilibrium in the one-stage auction $A(v^1, v^2..., v^n)$, such that $x_t = x$ for every $t > T_0$.*

The proof of Theorem A follows from combining the methods of proof of the following two weaker versions of it, Propositions 1 and 2. These versions are needed for later references. In Proposition 1 we prove our convergence result under the assumption that every player is using a learning scheme with a bounded recall, and in Proposition 2 we prove the same result under the assumption that every player is using a GFP learning scheme.

PROPOSITION 1. *Let $RA(v^1, v^2, ..., v^n)$ be a repeated first-price auction. Assume every player is using a learning scheme with a bounded recall, along with the tie-breaking rules TB1 and TB2. Then there exist a time T_0 and a strategy profile $x \in S$ which is in equilibrium in the one-stage auction $A(v^1, v^2..., v^n)$, such that $x_t = x$ for every $t > T_0$.*

PROPOSITION 2. *Let $RA(v^1, v^2, ..., v^n)$ be a repeated first-price auction. Assume every player is using a GFP learning scheme, along with the tie-breaking rules TB1 and TB2.Then there exist a time T_0 and a strategy profile $x \in S$ which is in equilibrium in the one-stage auction $A(v^1, v^2..., v^n)$, such that $x_t = x$ for every $t > T_0$.*

We end the section with a remark concerning a possible, seemingly shorter proof of Proposition 2.

A path $(x_1, x_2, ...)$, in S is a *better-reply* path if for every $t \geqslant 1$ for which x_t is not in equilibrium, $x_{t+1} \neq x_t$ and for every i for which $x^i_{t+1} \neq x^i_t$, Player i strongly prefers x^i_{t+1} to x^i_t when he believes that the next move of all other players is x^{-i}_t. Monderer and Sela [20] proved that if all players use a GFP learning schemeand apply the tie breaking rule TB1, then eliminating all successive repetitions from the path generated by the players yields a better-reply path. They deduce that in a game that does not have better reply cycles, the path generated by players that use GFP learning schemes and apply the tie-breaking rule TB1, must stabilize on equilibrium. One may think that an auction with commonly known types has this non-cycling property. This would have provided a very short proof of Proposition 2. The next example shows, however, that this is not the case.[9]

[9] Monderer and Sela [20] conjecture a stronger form of the above mentioned theorem: If the game does not have better reply cycles of length greater than two, then the path generated by players that use GFP learning schemes and apply the tie-breaking rule TB1, stabilizes on equilibrium. As our example shows a cycle of length three, even if the stronger version holds, it does not apply here.

EXAMPLE 1. Consider two players with $v^1 = v^2 = 7$. A better reply cycle is:

$$(5, 2), (2, 5), (6, 2), (5, 2), (2, 5), (6, 2), \ldots .$$

5. REMOVING THE TIE-BREAKING RULE TB1: BELIEF CONVERGENCE

We first show by an example that Proposition 2 does not hold without the tie-breaking rule TB1.

EXAMPLE 2. There are two players. $v^1 = 9$, $v^2 = 5$. Both players use a FP learning scheme with $e^1 = e^2 = 1$. The players may generate the following path:

$$(5, 4), (5, 4), (5, 1), (5, 4), (5, 4), (5, 1), \ldots .$$

In this case, the path generated by the players does not stabilize. However, we show below that the corresponding belief path does stabilize.

For $t > \max\{e^1, e^2\}$, let

$$(p_t, q_t) = (B^2(h_{t-1}), B^1(h_{t-1}))$$

be the sequenceof beliefs. This sequence is converging to $(p, q) \in \Delta(S^1) \times \Delta(S^2)$, where $p(5) = 1$, $q(4) = \frac{2}{3}$, and $q(1) = \frac{1}{3}$. It is easily verified that (p, q) is a mixed-action equilibrium in the one-shot auction.[10] Note, however, that the players in Example 2 may generate a nonconverging belief sequence. For example, they may generate the path

$$(5, 4), (5, 4), (5, x_3^2), (5, 4), (5, 4), (5, x_6^2), \ldots, \tag{5.1}$$

where x_{3k}^2 is an arbitrary integer in $\{1, 2\}$. Although the belief sequence generated by the path in (5.1) does not necessarily converge, it approaches equilibrium in the sense of Monderer and Shapley [21]: A sequence $((p_t, q_t))_{t=1}^{\infty}$ in $\Delta(S^1) \times \Delta(S^2)$ *is approaching equilibrium* if for every $\varepsilon > 0$ there exists $T \geqslant 1$, such that (p_t, q_t) is an ε-equilibrium for every $t \geqslant T$.[11]

[10] Actually, it is well known that in a 2-person game in which each player uses a FP learning scheme, if the sequence of beliefs converges, then the limit point must be a mixed-action equilibrium.

[11] Equivalently, for every $\varepsilon > 0$ there exists $T \geqslant 1$, such that for every $t \geqslant T$, the Euclidean distance between (p_t, q_t) and the mixed-action equilibrium set is smaller than ε. Note that all famous convergence theorems for fictitious play (e.g., Robinson [25] (zero-sum games), and Miyasawa [19] (2×2 games)) prove that the belief sequence is approaching equilibrium and not necessarily converging to equilibrium.

THEOREM B. *Let $RA(v^1, v^2)$ be a repeated first-price auction with two players. Assume every player is using a GFP learning scheme, along with the tie-breaking rule TB2. Thenthe belief sequence generated by the players is approaching equilibrium in the one-stage auction $A(v^1, v^2)$.*

The proof of Theorem B is given in Section 8. In the next example we assume that both players use a FP learning schemewith bounded recall. It is shown that the pathof actions that is generated by the players does not stabilize and the belief sequence does not approach equilibrium.

EXAMPLE 3 (Samuelson). There are two players, $v^1=9$, $v^2=5$. Both players use a learning scheme with a recall of size 1. The players may generate the following path (cycle):

$$(5, 1), (2, 2), (3, 3), (4, 4), (5, 4), (5, 1), \ldots.$$

The belief sequence is converging to $(p, q) \in \Delta(S^1) \times \Delta(S^2)$, where $p(5) = \frac{2}{5}$, $p(4) = p(3) = p(2) = \frac{1}{5}$, $q(4) = \frac{2}{5}$, and $q(3) = q(2) = q(1) = \frac{1}{5}$. As $q(1) > 0$, and 1 is not a best-response to p, then (p, q) is not in equilibrium.

When we deal with $n \geqslant 3$ players, any limit point of the belief sequence belongs to the set $\times_{i \in N} \Delta(S^{-i})$, and therefore, it is meaningless to discuss approaching an equilibrium of the belief sequence. However, if every player is using a FP learning scheme, we can define $p_t^i \in \Delta(S^i)$ as the empirical distribution of Player i's actions up to time t and ask whether the sequence $(p_t^1, p_t^2, \ldots, p_t^n)$ is approaching equilibrium. The next example shows that this is not necessarily the case.

EXAMPLE 4. Let $v^1 = 100$, $v^2 = v^3 = 98$. Assume the players use FP learning schemes. They may generate the path:

$$(98, 97, 1), (98, 1, 97), (98, 97, 1), (98, 1, 97), \ldots.$$

The individual empirical distribution vector is converging to (p^1, p^2, p^3), where $p^1(98) = 1$, and for $i = 2, 3$, $p^i(97) = p^i(1) = \frac{1}{2}$. It is easily verified that for Player 1, the bid 98 is not a best-response to (p^2, p^3). Therefore (p^1, p^2, p^3) is not in equilibrium.

6. OTHER LEARNING SCHEMES

In this section we discuss belief-based learning schemes, which seem natural in the context of repeated auction, but which are not covered by Theorem A. The original definition of fictitious play was given for 2-person games. One possible generalization to more than 2 players is the one given in Section 5. One may consider another possible generalization as in Monderer and Shapley [21]. In this version, we say that Player i uses the *individual fictitious play* (*IFP*) *learning scheme* if he acts myopically and at every stage t he believes that each of his opponents makes an independent decision and that for every $j \neq i$, Player j's next choice is distributed according to j's empirical distribution up to stage $t-1$. One can similarly define generalized IFP learning schemes and individual learning schemes with bounded recall. Except for 2-person games when the FP and IFP learning schemes coincide, we do not know whether any of our convergence results holds for the IFP learning schemes.

Next, consider a learning scheme in which Player i believes that the next maximal bid will be the average of all previous maximal bids.[12] Since bids must be integers and the average maximal bid is not necessarily an integer, we slightly modify the model. If the average maximal bid of all other players is a, where $l < a < l+1$, l a positive integer, then Player i assigns probability $l+1-a$ to l and $a-l$ to $l+1$. We call such a learning scheme an *average maximal bid* (*AMB*) learning scheme. Hon-Snir [7] proved (for the discrete first-price auction discussed in this paper) that when all players use an AMB learning scheme and apply the tie breaking rules TB1 and TB2, then the generated path of action profiles stabilizes on equilibrium.

7. ADDITIONAL REMARKS: FUTURE RESEARCH

Domination. Hon-Snir [7] proved that the discrete first-price auction game discussed in this paper is weak dominance solvable in the sense of Moulin [23]. That is, it is solved by successive elimination of *all* weakly dominated strategies, in the sense that every strategy profile in the Cartesian product of the sets of strategies that survive the elimination process, is in equilibrium. It is tempting to make the conjecture that Theorem A remains true for such games. Technically, Theorem A assumes TB2 which has no meaning in general games. Consider, however, the game derived from the first-price auction game by eliminating for each player, all bids which are greater than or equal to his valuation. From Theorem A we can conclude that for every such quasi first-price auction game, if the players use one of the learning schemes discussed in this paper, along with TB1, then the path of actions eventually stabilizes on equilibrium. But as the following example shows, even Proposition 1 does not hold for general weak dominance solvable games.

[12] The belief function in this scheme takes deterministic values. Thorlund-Peterson [29] discusses such learning scheme applied to the Cournot game.

EXAMPLE 5. Consider the following two-person game in which Player 1 chooses a row, and Player 2 chooses a column:

$$\begin{array}{cc} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} 0,1 & 0,1 & 1,0 \\ 1,0 & 2,2 & 1,0 \\ 1,0 & 0,1 & 0,1 \end{pmatrix}. \end{array}$$

This game is weak dominance solvable and the elimination process leads to the outcome *bb*. Assume each player is using a learning scheme with a recall of size 1, along with the tie breaking rule TB1. If the initial move is *aa*, the path of actions generated by the players may follow the cycle *aa*, *ca*, *cc*, *ac*, *aa*.

Consider a game which is solvable by successive elimination of strongly dominated strategies. One can deduce from Milgrom and Roberts [18], that if every player is using a GFP learning scheme with the tie-breaking rule TB1, then the path of action profiles generated by the players stabilizes on equilibrium. It is easy to show that this result holds also when each player is using either a learning scheme with bounded recall, or a GFP learning scheme. That is, Theorem A is valid for such games. Obviously a first-price auction game is not strong dominance solvable. One may expect some intermediate dominance solvability property in between strong and weak, which is satisfied by first-price auction games and implies Theorem A.[13] Such a natural property is the invariance under order of elimination. However, the game in Example 5 has this property. Rochet [26] characterized a subclass of games that satisfy this property. In these games, if one player is indifferent to two action profiles so are the other players. As noted by Marx and Swinkels [10], discrete first-price auction games do not satisfy Rochet's condition. Nevertheless it is interesting to know that Rochet's condition does not imply our learning result, as is shown in the following example.[14]

[13] Moulin [23] proved that if a game is weak dominance solvable and in addition has the property that the (pure strategy) best response correspondences are singled valued, then if all players use a bounded recall learning scheme with a recall of size 1, the process stabilizes on equilibrium. We do not know whether the same result holds when the players have bigger recalls, though we conjecture it does not. However, first-price auction games do not satisfy Moulin's condition.

[14] Marx and Swinkels [10] generalized Rochet's theorem by proving it under a weaker requirement which they call TDI (*transference of decisionmaker indifference*). As they noted, generically discrete first-price auction games satisfy TDI, but Example 6 shows TDI is not sufficient for our learning result. It is interesting to note that the first-price auction games discussed in this paper do not satisfy TDI.

EXAMPLE 6. Consider the following two-person game in which Player 1 chooses a row, and Player 2 chooses a column.

$$\begin{array}{cc} & \begin{array}{ccc} a & b & c \end{array} \\ \begin{array}{c} a \\ b \\ c \end{array} & \begin{pmatrix} 2,6 & 6,4 & 6,4 \\ 5,2 & 6,4 & 6,4 \\ 0,0 & 8,1 & 5,3 \end{pmatrix} \end{array}$$

This game is weak dominance solvable and the elimination process leads to the outcome *bc*. Assume each player is using a learning scheme with a recall of size 1, along with the tie breaking rule TB1. If the initial move is *aa*, the path of actions generated by the players may follow the cycle *aa*, *ba*, *bb*, *cb*, *cc*, *ac*, *aa*.

To summarize, while dominance solvability seems to play a crucial role in Theorem A, it seems that first-price auction games have a special additional structure which is not easily identified and which forces convergence.[15]

Imperfect monitoring. In the models discussed in this paper we assume "perfect monitoring." That is, at every stage t, every player i knows the full history of bids up to time $t-1$, or at least he knows the full history of the last m^i bids. In the context of auctions it is reasonable to assume that the players are informed only about the winning bids. It seems to us that analyzing repeated auctions when the players are using belief based learning schemes with such imperfect monitoring will contribute to auction theory.

Reinforcement learning. In the theory of reinforcement learning,[16] players do not form beliefs about the other players' next move. They are assumed to use a mixed action at each stage, where the probability assigned by this mixed action to a pure bid positively depends on the success of this bid in the past. The many ways in which these probabilities can be updated give rise to a variety of reinforcement strategies. It seems natural to analyze repeated auctions with reinforcement players.

Varying types. Consider a repeated auction in which the players' types vary stochastically with time. If the distribution of the random type vectors does not depend on time, then we actually deal with a repeated Bayesian game. If all the players are using learning schemes, then they generate a

[15] Note that such games do not satisfy any of the known conditions for convergence of the FP process which do not involve domination properties. That is, they are not zero-sum games (Robinson [25]), they are not supermodular games (Krishna [8]), and they are not weighted potential games (Monderer and Shapley [21]).

[16] See, e.g., Roth and Erev [27].

stochastic process in S. Hon-Snir [7] partially analyzed the stochastic path generated by fictitious players in a model where players' types are determined at each stage by the same i.i.d. random variables, each of them is uniformly distributed on $\{1, 2, ..., \bar{V}\}$. She shows that if the number of possible types for each player does not exceed seven (i.e., $\bar{V} \leqslant 7$), then with probability one, for a sufficiently late stage the players' behavior is in equilibrium in the one-stage Bayesian game in which the (common) distribution of each type is commonly known. She used computer simulation to analyze the model with more than seven possible types. It seems that the result continues to hold, although no analytical proof is given.

Removing the TB2 assumption. We conjecture that all our theorems hold without the TB2 assumption, but it increases the size of the proofs significantly.[17] Since this is a natural assumption we do not actually prove this conjecture.

8. PROOFS

Proof of Theorem A. We first prove two propositions. These propositions are needed for later reference. The methods of proof of the propositions can be combined in an obvious manner to generate a proof of Theorem A.

PROPOSITION 1. *Let $RA(v^1, v^2, ..., v^n)$ be a repeated first-price auction. Assume every player is using a learning scheme with a bounded recall, along with the tie-breaking rules TB1 and TB2. Then there exist a time T_0 and a strategy profile $x \in S$ which is in equilibrium in the one-stage auction $A(v^1, v^2 ..., v^n)$, such that $x_t = x$ for every $t > T_0$.*

Proof of Proposition 1. Denote the recall size of Player i by m^i. Assume without loss of generality that Nature chooses the types in a nonincreasing order. That is, $v^1 \geqslant v^2 \geqslant \cdots \geqslant v^n$. The proposition obviously holds when $v^2 = 1$. We therefore proceed to prove it under the assumption that $v^2 > 1$. Let $h = (x_1, x_2, ...)$ be the path generated by the players. We will need also the following notations: Let $e = \max_{j \in N} e^j$, let M^i be the set of all players j for which $v^j = v^i$, and let $y_t(i) = \max_{j \in M^i} x_t^j$. For $t > e^j$, let $p_t^j \in \Delta(S^{-j})$ be the belief of j about the tth actions of the other players. That is, $p_t^j = B^j(x_1, ..., x_{t-1})$. Let $p_t^j[b]$ denote the p_t^j-probability that the maximal bid of all other players is b, and let

$$q_t^j[b] = \sum_{\varnothing \neq M \subseteq N \setminus \{j\}} \frac{1}{|M| + 1} p_t^j(x^i = b \text{ for } i \in M;\ x^i < b \text{ for } i \notin M),$$

[17] It also enlarges the set of possible limit equilibria.

where $|M|$ denotes the number of players in M. Note that in the non-purified model $q_t^j[b]$ is the probability of winning if the maximal bid of all other players is b and j's bid is also b. Therefore, if j bids b at stage t, then according to his belief, his expected utility is

$$E_t^j(b) = (v^j - b)(p_t^j[1] + \dots + p_t^j[b-1] + q_t^j[b]).$$

Note further that

$$\frac{1}{n}\, p_t^j[b] \leqslant q_t^j[b] \leqslant \frac{1}{2}\, p_t^j[b]. \tag{8.1}$$

We need the following claim.

CLAIM 1. *For every $j \in N$ and for every $t > e$, $x_t^j \leqslant v^2$.*

Proof of Claim 1. If $v^1 = v^2$, then $v^j \leqslant v^2$ for every player j and therefore the assertion follows from TB2. If $v^1 > v^2$, then all players in $N \backslash \{1\}$ bid less than v^2 by TB2. Therefore, for $t > e \geqslant e^1$, a best response of Player 1 cannot exceed v^2. ▮

We proceed to show that there exists T_0 such that for every $j \in M^1$ and for every $t > T_0$, $x_t^j \geqslant v^2 - 2$ and in addition, if M^1 is a singleton, then $y_t(2) \geqslant v^2 - 2$ for each such t.[18]

This is obvious if $v^2 < 4$, thus we proceed to prove it, assuming that $v^2 \geqslant 4$. We prove by induction on $1 \leqslant k \leqslant v^2 - 3$ that there exists T_k such that $x_t^j \geqslant k + 1$ for every $j \in M^1$ and $t > T_k$ and that if M^1 is a singleton then $y_t(2) \geqslant k + 1$ for each such t. Additional two claims are needed:

CLAIM 2. *Let $1 \leqslant k \leqslant k + 1 \leqslant v^j - 1$. If at time $t > e$, Player j weakly prefers k to $k + 1$, then*

$$p_t^j[1] + \cdots + p_t^j[k-1] \geqslant \frac{v^j - k - 2}{2}\, p_t^j[k], \tag{8.2}$$

where the left-hand side of (8.2) *equals zero when $k = 1$.*

Proof of Claim 2. As j weakly prefers k to $k + 1$, $E_t^j(k) \geqslant E_t^j(k+1)$. Therefore

$$(v^j - k)(p_t^j[1] + \cdots + q_t^j[k]) \geqslant (v^j - k - 1)(p_t^j[1] + \cdots + q_t^j[k+1]).$$

[18] Note that we do not claim that all players in M^2 make a bid greater than $v^2 - 3$. For example, if $v^1 = 9$, $v^2 = v^3 = 5$, and the players have recall of size 1, then M^1 is a singleton, 2 and 3 belong to M^2, and the players may generate the path: (5, 4, 1), (5, 4, 1), ...

As $q_t^j[k+1] \geqslant 0$ and by (8.1), $q_t^j[k] \leqslant \frac{1}{2} p_t^j[k]$, the last inequality yields

$$p_t^j[1] + \cdots + p_t^j[k-1] \geqslant \left((v^j - k - 1) - \frac{v^j - k}{2} \right) p_t^j[k].$$

Hence, (8.2) is obtained by manipulating the right-hand side of the previous inequality. ▌

The proof of the next claim is obvious.

CLAIM 3. *Let* $1 \leqslant k \leqslant v^j - 2$. *If* j *weakly prefers* k *to* $v^j - 1$ *at* $t > e$, *then*

$$(v^j - k)(p_t^j[1] + \cdots + q_t^j[k]) \geqslant p_t^j[1] + \cdots + q_t^j[v^j - 1].$$

We now return to the main proof.

$k = 1$. Let $j \in M^1$ (i.e., $v^j = v^1$). We show that j does not bid 1 for $t > e$. Indeed, assume in negation that $x_t^j = 1$ for such t. In particular, j weakly prefers 1 to 2 at t. Hence, by Claim 2,

$$0 \geqslant \frac{v^1 - 3}{2} \; p_t^j[1].$$

As $v^1 > 3$, $p_t^j[1] = 0$. This implies by (8.1) that $q_t^j[1] = 0$. As j weakly prefers 1 to $v^1 - 1$, Claim 3 yields

$$0 = (v^1 - 1)\, q_t^j[1] \geqslant p_t^j[1] + \cdots + p_t^j[v^1 - 2] + q_t^j[v^1 - 1].$$

This implies that one of the other players, say player i, chose $x_s^i \geqslant v^j$ for some $t - m^j \leqslant s \leqslant t - 1$, contradicting TB2.

Assume now that M^1 is a singleton; that is, $v^1 > v^2$. Let $i_t \in M^2$ be a player with $x_t^{i_t} = y_t(2)$. We show that i_t does not bid 1 for $t > 2e$. If $x_t^{i_t} = 1$ then $x_t^i = 1$ for every $i \in M^2$. Let $i \in M^2$. As in the previous paragraph, the fact that i weakly prefers 1 to 2 implies $p_t^i[1] = 0$. The fact that i weakly prefers 1 to $v^2 - 1$ implies

$$0 = (v^i - 1)\, q_t^j[1] \geqslant p_t^i[1] + \cdots + p_t^i[v^2 - 2] + q_t^i[v^2 - 1].$$

Therefore, $x_s^1 = v^2$ for every $t - m^i \leqslant s \leqslant t - 1$. As by TB2 $x_{t-1}^i < v^2$, x_{t-1}^i is a best response to Player i's belief at time t and, therefore, by TB1, $x_t^i = x_{t-1}^i$. Hence, $x_{t-1}^i = 1$. At time $t - 1$, Player i bids 1 when he observes a history in which the maximal bid is v^2 for $m^i - 1$ times and the maximal bid is greater than 1 (because Player 1 bids more than 1 for $s > e$) in the first stage of this history; thus $x_{t-2}^i < v^2$ is a best response to the belief generated by this history, and by TB1, $x^i_{t-2}=1$. Continuing recursively, we show that for every $i\in M^2$, Player i bids 1 in the stages s, $t-m^1-1\leqslant s\leqslant t-2$. Therefore, at time $t-1$ Player 1 plays v^2 when he observes a history in which the maximal bid does not exceed v^2-2 (because every player $i\in M^2$ plays 1 in this history, and by TB2 any other player bids at most $v^3-1<v^2-1$). This is a contradiction because bidding v^2-1 gives a higher expected payoff than bidding v^2 versus such a belief. Assume the assertion holds for $k-1$, $2\leqslant k\leqslant v^2-3$, we now prove it for k with $T_k=T_{k-1}+2e$.

Let $j\in M^1$. Assume j bids k at some $t>T_k$. In particular j weakly prefers k to $k+1$. Therefore, by Claim 2, (8.2) holds. By the induction hypothesis the probability that the maximal bid of all other players is less or equal to $k-1$ equals zero, hence the left-hand side of (8.2) is zero and, therefore,

$$0\geqslant\frac{v^j-k-2}{2}\,p^j_t[k].$$

Since $v^j-k-2>0$, this yields $p^j_t[k]=0$. Since j weakly prefers k to v^j-1, we get, by Claim 3,

$$0=(v^j-k)\,q^j_t[k]\geqslant p^j_t[1]+\cdots+q^j_t[v^j-1].$$

Hence, there exists a player i who bid at least v^j along the last m^j moves, contradicting TB2. Assume now that M^1 is a singleton. Let $i_t\in M^2$ be a player with $x^{i_t}_t=y_2(t)$. Since i_t bids k at time t, then for every $i\in M^2$, $x^i_t\leqslant k$. Let $i\in M^2$, and denote x^i_t by τ. As i weakly prefers τ to $\tau+1$ we get (as before) that $p^i_t[\tau]=0$. Since i weakly prefers τ to v^2-1, we get, as before, that $p^i_t[1]+\cdots+q^i_t[v^2-1]=0$. Therefore, Player 1 played v^2 in the last m^i moves. By TB1, for every $i\in M^2$, $x^i_{t-1}=x^i_t$. This implies, as in the proof for $k=1$ that at time $t-1$ Player 1 played v^2 when he believed that with probability one the maximal bid did not exceed v^2-2. A contradiction.

We are now able to prove convergence and to characterize the limit point of the process.

Case 1 ($|M^1|=1$). Let $T_0>e$ be an integer such that for $s>T_0$, $x^1_s\geqslant v^2-2$ and $y_s(2)\geqslant v^2-2$. We show that for $t>T_0+e$, $x^1_t\geqslant v^2-1$. Assume in negation that Player 1 bids v^2-2 at such t. As he weakly prefers this bid to v^2-1, we get from Claim 2 and from the above property of T_0 that

$$0\geqslant\frac{v^1-v^2}{2}\,p^1_t[v^2-2]$$

and, hence, $p_t^1[v^2-2]=0$. Since Player 1 weakly prefers v^2-2 to v^1-1,

$$0=(v^1-v^2+2)\,q_t^1[v^2-2]\geqslant p_t^1[v^2-2]+\cdots+q_t^1[v^1-1].$$

Therefore one of the other players bid v^1 or more at one time along the last m^1 stages, contradicting TB2.

Case 1.1 ($v^1>v^2+1$). We show that there exists $\bar{T}$ such that for every $t>\bar{T}$, $x_t=x$, where x is an equilibrium satisfying $x^1=v^2$, there exists $i\in M^2$ such that $x^i=v^2-1$ for every player j with $v^j>1$, $x^j\leqslant v^j-1$, and $x^j=1$ if $v^j=1$.

Assume that for some stage $T^*>T_0+2+2e$, Player 1 bids v^2-1. Then, as was shown above in Case 1, right after that all players in M^2 observe an history in which the maximal bid in each step was either v^2-1 or v^2, and they assign a positive probability to the maximal bid being v^2-1. Therefore, they bid v^2-1. That is, $x^i_{T^*+1}=v^2-1$ for every $i\in M^2$. Therefore, for every $t>T^*+1$, every player in M^2 assigns a probability 1 to the maximal bid belonging to $\{v^2-1, v^2\}$, and thus, by TB1, $x_t^i=v^2-1$ for every such t. Therefore, for $t>T^*+m^1$, Player 1 observe an history of m^1 times in which the maximal bid was v^2-1. As $v^1>v^2+1$, Player 1 bids v^2. So, for $t>T^*+m^1$, $x_t=x$, where $x^1=v^2$, $x^i=v^2-1$ for every $i\in M^2$, $x^j\leqslant v^j-1$ for every player j with $v^j>1$, and $x^j=1$ when $v^j=1$.

Assume that for $t>\bar{T}=T_0+2+2e$, Player 1 bids only v^2. Then for every $t>\bar{T}+e+1$, $x_t=x$, where $x^1=v^2$ and for every $j\neq 1$, $x^j=x^j_{\bar{T}+e}$, and for one of the players $i\in M^2$, $x^i=v^2-1$ necessarily (since otherwise Player 1 would switch to v^2-1).

Case 1.2 ($v^1=v^2+1$). Assume that for $t>\bar{T}=T_0+2+2e$, Player 1 bids only v^2. Then we get the same convergence result as in Case 1.1. If, however, for some stage $T^*>T_0+2+2e$, Player 1 bids v^2-1, then as in Case 1.1, $x_t^i=v^2-1$ for every $i\in M^2$ and for every $t\geqslant T^*+1$. Therefore, for every $t>T^*+m^1$ Player 1 observes a history in which the maximal bid is constantly v^2-1. Unlike Case 1.1, this does not mean that 1 bids v^2 at stage T^*+m^1+1, because it may be that v^2 is not his unique best response to such a history (if $|M^2|=1$, then v^2-1 is also a best reply). However, for every $t\geqslant T^*+m^1+1$, $x_t=x$, where $x^i=v^2-1$ for every $i\in M^2$, $x^j\leqslant v^j-1$ for every j with $v^j>1$, and $x^1=x^1_{T^*+m^1+1}$, where $x^1_{T^*+m^1+1}\in\{v^2-1, v^2\}$. Moreover, if $|M^2|>1$, then $x^1=v^2$.

Case 2 ($|M^1|>1$).

Case 2.1 ($|M^1|>2$). In this case we show that all players in M^1 bid v^1-1 after sufficiently large stage. That is, the process stabilizes at x, where $x^j=v^1-1$ for every $j\in M^1$ and $x^i\leqslant v^i-1$ for every player i. Note that $v^1=v^2$. Hence, there exists T^* such that for $t>T^*$, each player in M^1 bids v^1-2 or v^1-1. Therefore, for every $t>T^*+e$, every player j in M^1

assigns a probability 1 to the maximal bid in $\{v^1-2, v^1-1\}$.We show that j strictly prefers v^1-1 to v^1-2. Indeed, assume j assigns a probability p to the maximal bid being v^1-2. If he bids v^1-2 his expected value is $E_2=2q_t^j[v^1-2]$. If he bids v^1-1, his expected value is $E_1=p_t^j[v^1-2]+q_t^j[v^1-1]$. If $p<1$, then $q_t^j[v^1-1]>0$ and, therefore,

$$\begin{aligned}E_1 &= p_t^j[v^1-2]+q_t^j[v^1-1]\\ &\geqslant 2q_t^j[v^1-2]+q_t^j[v^1-1]>2q_t^j[v^1-2]=E_2.\end{aligned}$$

If $p=1$, then Player j observes an history of m^j times in which the maximal bid of all other players was v^2-2 and in which all other players in M^1 bid v^2-2. Because there are at least two other players in M^1, $q_t^j[v^1-2]\leqslant \frac{1}{3}p_t^j[v^1-2]$. Therefore $E_2\leqslant\frac{2}{3}E_1<E_1$.

Case 2.2 ($|M^1|=2$). In this case both players in M^1 bid either v^1-1 or v^1-2 for a sufficiently large stage. If one of them bids v^1-1once, he will continue this bid forever, because of TB1. Therefore, eventually the other player switches to v^1-1 too. So, the process stabilizes at $x^1=x^2=v^1-1$, and $x^j\leqslant v^j-1$ for every player j. It may be, however, that both players play v^1-2 forever, provided that $v^3<v^2-1$. If $v^3=v^2-1$, then necessarily players 1 and 2 bid v^1-1 from a certain point on, because otherwise the players in M^3 switch to v^2-2 and there after make v^1-1 a strictly best reply for the players in M^1. ▮

PROPOSITION 2. *Let $RA(v^1, v^2, ..., v^n)$ be a repeated first-price auction. Assume every player is using a GFP learning scheme, along with the tie-breaking rules TB1 and TB2. Then there exist a time T_0 and a strategy profile $x\in S$ which is in equilibrium in the one-stage auction $A(v^1, v^2..., v^n)$, such that $x_t=x$ for every $t>T_0$.*

Proof of Proposition 2. Assume without loss of generality that Nature chooses the types in a nonincreasing order. That is, $v^1\geqslant v^2\geqslant\cdots\geqslant v^n$. The proposition holds obviously when $v^2=1$. We, therefore, proceed to prove it under the assumption that $v^2>1$. Let $h=(x_1, x_2, ...)$ be the path generated by the players. Let $e=\max_{j\in N}e^j$, let M^i be the set of all players j for which $v^j=v^i$, and let $y_t(i)=\max_{j\in M^i}x_t^j$. Using the rest of the notations established in the proof of Proposition 1, it can be seen that Claims 1, 2, and 3 continue to hold. We proceed to prove another claim.

CLAIM 4. *Every $j\in M^1$ makes a bid in $\{1, 2, ..., v^2-3\}$ only finitely many times. Moreover, if M^1 is a singleton and*

A1. *Player 1 makes a bid in $\{1, 2, ..., v^2-1\}$ infinitely many times,*

then every $i\in M^2$ makes a bid in $\{1, 2, ..., v^2-3\}$ only finitely many times.

Proof of Claim 4. This claim is obvious if $v^2 < 4$; thus we proceed to prove it assuming that $v^2 \geqslant 4$. We prove by induction on $1 \leqslant k \leqslant v^2 - 3$, that every $j \in M^1$ makes a bid in $\{1, ..., k\}$ only finitely many times and that, in addition, if M^1 is a singleton and Assumption A1 holds, then every $i \in M^2$ makes a bid in $\{1, ..., k\}$ only finitely many times.

$k = 1$. Let $j \in M^1$ (i.e., $v^j = v^1$). We show that for $t > e$, Player j does not bid 1. Indeed, assume in negation that $x_t^j = 1$ for such t. In particular, j weakly prefers 1 to 2 at t. Therefore, by Claim 2, $p_t^j[1] = 0$. Hence By Claim 3 (as Player j weakly prefers 1 to $v^1 - 1$),

$$0 = p_t^j[1] + \cdots + p_t^j[v^1 - 2] + q_t^j[v^1 - 1].$$

This implies that for some $w \geqslant v^1$, $p_t^j[w] > 0$. Therefore, by AD2, there exists a player i such that for some $1 \leqslant s \leqslant t-1$, $x_s^i = w \geqslant v^1 \geqslant v^i$, in contradiction to TB2.

Assume M^1 is a singleton, i.e., $v^1 > v^2$, and that A1 holds. Let $i \in M^2$. We show that Player i bids 1 only finitely many times. Assume in negation that i bids 1 infinitely many times, at times $e < t_1 < t_2 < \cdots$. At time t_l Player i weakly prefers 1 to 2; thus by Claim 2, $p_{t_l}^i[1] = 0$. As Player i weakly prefers 1 to $v^2 - 1$, Claim 3 yields

$$0 = p_{t_l}^i[1] + \cdots + p_{t_l}^i[v^2 - 2] + q_{t_l}^i[v^2 - 1].$$

By Claim 1, the last equality yields $x_s^1 = v^2$ for every $1 \leqslant s \leqslant t_l - 1$. As $\lim_{l \to \infty} t_l = \infty$, $x_s^1 = v^2$ for every $s \geqslant 1$, contradicting A1.

Assume the assertion holds for $k-1$, $2 \leqslant k \leqslant v^2 - 3$, we now prove it for k. Let $j \in M^1$. Assume in negation that Player j bids k infinitely many times, at times $e < t_1 < t_2 \cdots$. When Player j bids k at t_l, he weakly prefers k to $k+1$. Therefore by Claim 2, for every $l \geqslant 1$

$$w_j(t_l - 1)(p_{t_l}^j[1] + \cdots + p_{t_l}^j[k-1]) \geqslant w_j(t_l - 1)\left(\frac{v^1 - k - 2}{2}\, p_{t_l}^j[k]\right).$$

By the induction hypothesis the left-hand side of the last inequality is bounded when l varies, say by M. Therefore, the right hand-side is also bounded by M. As j weakly prefers k to $v^1 - 1$ we get from Claim 3:

$$\begin{aligned} &w_j(t_l - 1)(v^1 - k)(p_{t_l}^j[1] + \cdots + q_{t_l}^j[k]) \\ &\qquad \geqslant w_j(t_l - 1)(p_{t_l}^j[1] + \cdots + q_{t_l}^j[v^1 - 1]). \end{aligned}$$

Since the left-hand side of this inequality is bounded, so is the right-hand side. This implies that the bids $\{1, ..., v^1 - 1\}$ were used only finitely many times, contradicting TB2. Assume now that M^1 is a singleton and A1 holds. Let $i \in M^2$. We show that Player i bids k only finitely many times. Assume in negation that Player i bids k infinitely many times, at times $e < t_1 < t_2 < \cdots$. At time t_l Player i weakly prefers k to $k+1$; thus we conclude, as in the first part of the kth step, that there exists M such that

$$w_i(t_l - 1)(p^i t_l[1] + \cdots + q^i_{t_l}[k]) \leqslant M \qquad \text{for every} \quad l \geqslant 1. \tag{8.3}$$

Since at stage t_l Player i weakly prefers k to $v^2 - 1$, Claim 3 and (8.3) yield for every $l \geqslant 1$

$$w_i(t_l - 1)(p^i_{t_l}[1] + \cdots + p^i_{t_l}[v^2 - 2] + q^i_{t_l}[v^2 - 1]) \leqslant M(v^2 - k).$$

The last inequality implies that there exists T_0 such that Player 1 bids v^2 for every $t \geqslant T_0$ in contradiction to A1. This completes the proof of Claim 4. ∎

We are now able to prove convergence and to characterize the limit point of the process.

Case 1 ($|M^1| = 1$). We need the following claim.

CLAIM 5. *Let M^1 be a singleton. Then Player* 1 *bids $v^2 - 2$ only finitely many times.*

Proof of Claim 5. Assume 1 bids $v^2 - 2$ infinitely many times, at times $e < t_1 < t_2 \cdots$. When 1 bids $v^2 - 2$ at t_l, he weakly prefers $v^2 - 2$ to $v^2 - 1$. Therefore, by Claim 2 for every $l \geqslant 1$

$$w_1(t_l - 1)(p^1_{t_l}[1] + \cdots + p^1_{t_l}[v^2 - 3]) \geqslant w_1(t_l - 1)\,\frac{v^1 - v^2}{2}\, p^1_{t_l}[v^2 - 2].$$

By Claim 4 the left-hand side of the last inequality is bounded when l varies, say by M. Therefore the right-hand side is also bounded by M. As 1 weakly prefers $v^2 - 2$ to $v^1 - 1$ we get from Claim 3

$$\begin{aligned} w_1(t_l - 1)(v^1 - v^2 + 2)(p^1_{t_l}[1] + \cdots + q^1_{t_l}[v^2 - 2] \\ \geqslant w_1(t_l - 1)(p^1_{t_l}[1] + \cdots + q^1_{t_l}[v^1 - 1]). \end{aligned}$$

Since the left-hand side of this inequality is bounded, so is the right-hand side. This implies that the bids $\{1, ..., v^1 - 1\}$ are used only finitely many times, contradicting TB2. ∎

Case 1.1 ($v^1 > v^2 + 1$) *or* ($v^1 = v^2 + 1$ *and M^2 is not a singleton*). We show that there exists $\overline{T}$ such that for every $t \geqslant \overline{T}$, $x_t = x$, where x is in equilibrium in the auction $A(v^1, v^2, ..., v^n)$ satisfying $x^1 = v^2$, there exists $i \in M^2$ with $x^i = v^2 - 1$, and $x^j \leqslant v^j - 1$ for every player j with $v^j > 1$.

Assume A1 is not satisfied. Then there exists $T_0 > e$ such that Player 1 bids v^2 for $t \geqslant T_0$. Therefore,[19] by TB1, for each player $j \neq 1$, $x_t^j = x_{T_0}^j$ for every $t \geqslant T_0$. If $x_{T_0}^i < v^2 - 1$ for every $i \in M^2$, then Player 1 eventually switches from v^2. Therefore, the process stabilizes at x with $x^1 = v^2$; there exists $i \in M^2$ with $x^i = v^2 - 1$ and $x^j \leqslant v^j - 1$ for every player j. Assume A1 holds, then by Claim 5, Claim 4, and Claim 1, there exists $T_0 > e$ such that for every $t \geqslant T_0$, Player 1 makes bids in $\{v^2 - 1, v^2\}$ and every player in M^2 makes bids in $\{v^2 - 2, v^2 - 1\}$. Since A1 holds, player 1 bids $v^2 - 1$ infinitely many times. Therefore, for sufficiently large t, for each $i \in M^2$, the conditional probability of the maximal bid of the other players being $v^2 - 1$, given that this maximal bid is less than v^2, is increasing to 1. Therefore, there exists a stage when all players in M^2 switch to $v^2 - 1$ and stay with this bid. Since $v^1 > v^2 + 1$, or $v^1 = v^2 + 1$ and M^2 is not a singleton, there exists a stage when Player 1 switches to v^2 and stay with this *bid*, which contradicts A1. Therefore, Assumption A1 cannot hold.

Case 1.2 ($v^1 = v^2 + 1$ *and* M^2 *is a singleton.*). If Assumption A1 does not hold, then we get the same convergence as in Case 1.1. If A1 does hold, then, as in Case 1.1, Player 2 bids $v^2 - 1$ for sufficiently large t. In contrast to Case 1.1, it is no longer true that this forces Player 1 to switch to v^2. Actually, if Player 2 made at any stage in the past a bid smaller than $v^2 - 1$, then Player 1 must eventually bid $v^2 - 1$ (because he uses a best-response). If Player 2 made only the bid $v^2 - 1$, then both v^2 and $v^2 - 1$ are best-response actions for Player 1. Because A1 and TB1 hold, then Player 1 uses $v^2 - 1$ for sufficiently large t. The process therefore stabilizes at x, with $x^1 = x^2 = v^2 - 1$ and $x^j \leqslant v^j - 1$ for every Player j with $v^j > 1$.

Case 2 ($|M^1| \geqslant 2$). $v^1 = v^2$. Thus by Claim 4, there exists T^* such that for $t > T^*$, each player in M^1 bids $v^1 - 2$ or $v^1 - 1$.

Claim 6. *Suppose* $|M^1| \geqslant 2$. *If there exists a player in* M^1 *who bids* $v^1 - 2$ *infinitely many times, then there exists* T_0 *such that for* $t > T_0$ *all players in* M^1 *bid* $v^1 - 2$.

Proof of Claim 6. Let $j \in M^1$ bids $v^1 - 2$ infinitely many times, at times $T^* < t_1 < t_2 < \cdots$. Then for every $l \geqslant 1$, $E_{t_l}^j(v^1 - 2) \geqslant E_{t_l}^j(v^1 - 1)$. Hence,

$$2(p_t^j[1] + \cdots + q_t^j[v^1 - 2]) \geqslant p_t^j[1] + \cdots + p_t^j[v^1 - 2] + q_t^j[v^1 - 1].$$

[19] Here we use the special structure of GFP learning schemes that implies that for $t \geqslant T_0$, the conditional probability measure, given that the maximal bid is smaller than v^2 remains fixed.

Therefore,

$$(p_{t_l}^j[1] + \cdots + p_{t_l}^j[v^1-3]) \geqslant p_{t_l}^j[v^1-2] - 2q_{t_l}^j[v^1-2] + q_{t_l}^j[v^1-1]. \quad (8.4)$$

By (8.1), $p_{t_l}^j[v^1-2] \geqslant 2q_{t_l}^j[v^1-2]$.Therefore,

$$(p_{t_l}^j[1] + \cdots + p_{t_l}^j[v^1-3]) \geqslant q_{t_l}^j[v^1-1]. \quad (8.5)$$

Multiplying both sides of (8.5) by $w^j(t_l)$ gives a bounded left-hand side (when l varies) and therefore, a bounded right-hand side. Thus, there exists T_0 such that $x_{t_l}^d = v^2 - 2$ for every $d \in M^1$, $d \neq j$. Let $d \in M^1$, $d \neq j$. Since d plays $v^2 - 2$ infinitely many times, we get, replacing j with d in the above calculation that all players in M^1 other than Player d play $v^1 - 2$ for sufficiently large t. Since Player d plays $v^2 - 2$ for sufficiently large t, all players in M^1 play $v^2 - 2$ for sufficiently large t. ∎

Case 2.1 ($(|M^1| > 2)$ *or* $(|M^1| = 2$ *and* $v^3 = v^1 - 1)$). In this case we show that all players in M^1 eventually bid $v^1 - 1$. That is, the process stabilizes at x, where $x^j = v^1 - 1$ for every $j \in M^1$, and $x^i \leqslant v^i - 1$ for every player i. Indeed, if our assertion does not hold, then, by Claim 6, all players in M^1 bid $v^2 - 2$ for sufficiently large t. If there are more than two players in M^1, or $v^3 = v^1 - 1$, then by Claim 6, for sufficiently large t each player j in M^1 believes that with a high probability the maximal bid of the other players is $v^2 - 2$ and that there are at least two other players who bid $v^2 - 2$. This forces Player j to switch to $v^2 - 1$. A contradiction.

Case 2.2 ($(|M^1| = 2)$ *and* $(v^3 < v^1 - 1)$). In this case, by what we have shown, the process stabilizes at some equilibrium x, of one of two possible forms: Either $x^1 = x^2 = v^1 - 1$ and $x^i \leqslant v^i - 1$ for every player i with $v^i > 1$, or $x^1 = x^2 = v^1 - 2$ and $x^i \leqslant v^i - 1$ for every player i with $v^i > 1$. ∎

Proof of Theorem B. Denote the belief sequence by $((p_t, q_t))_{t \geqslant 2}$. We use the following characterization for approaching equilibrium given in Monderer and Shapley [21]: $((p_t, q_t))_{t \geqslant 1}$ is approaching equilibrium if and only if every limit point of this sequence is in equilibrium. We also use Claims 1–6, which were proved without utilizing TB1. Let (p, q) be a limit point of the belief sequence.

Case 1.1 ($v^1 > v^2 + 1$). Assume A1 is not satisfied. Then there exists $T_0 > e$ such that Player 1 bids v^2 for $t \geqslant T_0$. Therefore, p is the probability measure concentrated on v^2. As for every $t \geqslant T_0$, v^2 is a best response to q_t, p is a best response to q. On the other hand, by TB2, q assigns a positive probability only to bids in $\{1, ..., v^2 - 1\}$ and each bid in this set is a best response to v^2. Therefore, q is a best response to p. Hence (p, q) is in equilibrium.

Assume A1 holds, then by Claim 5, Claim 4, and Claim 1 there exists $T_0 > e$ such that Player 1 makes bids in $\{v^2 - 1, v^2\}$ and Player 2 makes bids in $\{v^2 - 2, v^2 - 1\}$ for every $t \geqslant T_0$. Since A1 holds, Player 1 bids $v^2 - 1$ infinitely many times. Therefore, for sufficiently large t, Player 2's conditional probability of the bid of Player 1 being $v^2 - 1$, given that this bid is less than v^2, is increasing to 1. Therefore, there exists a stage in which Player 2 switches to $v^2 - 1$ and stays with this bid. Since $v^1 > v^2 + 1$, there exists a later stage at which Player 1 switches to v^2 and stays with this *bid*, in contradiction to A1. Therefore Assumption A1 cannot hold.

Case 1.2 ($v^1 = v^2 + 1$). If Assumption A1 does not hold, then we get the same convergence as in Case 1.1. If A1 does hold, then as in Case 1.1, Player 2 bids $v^2 - 1$ for sufficiently large t. Therefore, $q = \delta_{v^2-1}$, where for a set X and for $x \in X$, δ_x is the probability measure concentrated on x. In contrast to Case 1.1, it is no longer true that this forces Player 1 to switch to v^2. Actually, if Player 2 made at any past stage a bid smaller than $v^2 - 1$, then Player 1 must eventually bid $v^2 - 1$ (because he uses a best response). In this case $(p, q) = (\delta_{v^2-1}, \delta_{v^2-1})$ forms a pure action equilibrium. If Player 2 made only the bid $v^2 - 1$, then both v^2 and $v^2 - 1$ are best-response actions for Player 1. As p assigns a positive probability only to $v^2 - 1$ and v^2, and both these actions are best responses to $v^2 - 1$, p is a best response to q. As q is a best response to any mixture of $v^2 - 1$ and v^2, (p, q) is in equilibrium.

Case 2 ($v^1 = v^2$). Since $v^1 = v^2$, by Claim 4, there exists T^* such that for $t > T^*$, each player in M^1 bids $v^1 - 2$ or $v^1 - 1$.

If for every sufficiently late stage both players bid $v^1 - 1$, then $(p, q) = (\delta_{v^1-1}, \delta_{v^1-1})$, and therefore (p, q) is in equilibrium. If one of the players bids $v^2 - 2$ infinitely many times, then by Claim 6, (p, q) is the equilibrium $(\delta_{v^2-2}, \delta_{v^2-2})$. ∎

REFERENCES

1. O. Ashenfelter, How auctions work for wine and art, *J. Econ. Perspect.* **3** (1989), 23–36.
2. S. Bikhchandani, Reputation in repeated second-price auctions, *J. Econ. Theory* **46** (1988), 97–119.
3. P. C. Cramton, Money out of thin air: The nationwide narrowband PCS auction, *J. Econ. Manage. Strategy*, in press.
4. D. Fudenberg and D. Kreps, Learning mixed equilibria, *Games Econ. Behav.* **5** (1993), 320–367.
5. D. Fudenberg and D. Levine, Theory of learning in games, internet edition, September 24, 1996.
6. M. Harris and A. Raviv, Allocation mechanisms and the design of auctions, *Econometrica* **49** (1981), 1477–1499.
7. S. Hon-Snir, "Learning in Auctions," M.Sc. thesis, The Technion, Haifa, 1996. [Hebrew]

8. V. Krishna, Learning in games with strategic complementarities, mimeo, Harvard University, 1991.
9. J.-J. Laffont, Game theory and empirical economics: The case of auction data, *Europ. Econ. Rev.* **41** (1997), 1–35.
10. L. M. Marx and J. M. Swinkels, Order independence for iterated weak dominance, *Games Econ. Behav.* **18** (1997), 219–245.
11. R. P. McAfee, Mechanism design by competing sellers, *Econometrica* **61** (1993), 1281–1312.
12. R. P. McAfee and J. McMillan, Auctions and bidding, *J. Econ. Lit.* (1987), 699–738.
13. R. P. McAfee and D. Vincent, The declining price anomaly, *J. Econ. Theory* **60** (1993), 191–212.
14. J. McMillan, Selling spectrum rights, *J. Econ. Perspect.* **8** (1992), 145–162.
15. P. R. Milgrom, The economics of competitive bidding: A selective survey, *in* "Social Goals and Social Organizations" (L. Horwicz, D. Schmeidler, and H. Sonnenschein, Eds.), Cambridge Univ. Press, Cambridge, 1985.
16. P. R. Milgrom, Auction theory, *in* "Advances in Economic Theory: Fifth World Congress" (T. Bewley, Ed.), Cambridge Univ. Press, Cambridge, 1987.
17. P. R. Milgrom and R. J. Weber, A theory of auctions and competitive bidding, *Econometrica* **50** (1982), 1089–1122.
18. P. R. Milgrom and J. D. Roberts, Adaptive and sophisticated learning in normal form games, *Games Econ. Behav.* **3** (1991), 82–100.
19. K. Miyasawa, On the convergence of the learning process in a 2×2 nonzero-sum two person game, Economic Research Program, Princeton University, Research Memorandum No. 33, 1961.
20. D. Monderer and A. Sela, Fictitious play and no-cycling conditions, mimeo, 1993.
21. D. Monderer and L. S. Shapley, Fictitious play property for games with identical interests, *J. Econ. Theory* **1** (1996), 258–265.
22. R. B. Myerson, Optimal auction design, *Math. Oper. Res.* **6** (1981), 58–63.
23. H. Moulin, "Game Theory for the Social Sciences," 2nd ed., New York Univ. Press, New York, 1986.
24. J. G. Riley and W. F. Samuelson, Optimal auctions, *Amer. Econ. Rev.* **71** (1981), 381–392.
25. J. Robinson, An iterative method of solving a game, *Ann. Math.* **54** (1951), 296–301.
26. J. C. Rochet, Selection of an unique equilibrium value for extensive games with perfect information, D.P. Ceremade, Université Paris IX, 1980.
27. A. E. Roth and I. Erev, Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term, *Games Econ. Behav.* **8** (1995), 164–212.
28. R. M. Stark and M. H. Rothkopf, Competitive bidding: A comprehensive bibliography, *Oper. Res.* **27** (1979), 364–390.
29. L. Thorlund-Petersen, Iterative computation of cournot equilibrium, *Games Econ. Behav.* **2** (1990), 61–95.
30. W. Vickrey, Counterspeculation, auctions, and sealed tenders, *J. Finance* **16** (1961), 8–37.
31. R. J. Weber, Multiple-object auctions, *in* "Auctions, Bidding, and Contracting" (R. Engelbrecht-Wiggans, M. Shubic, and J. Stark, Eds.), pp. 165–191, New York Univ. Press, New York, 1983.
32. R. Wilson, Strategic analysis of auctions, *in* "Handbook of Game Theory, Vol. 1" (R. J. Aumann and S. Hart, Eds.), Elsevier Science, New York, 1992.
33. E. Wolfstetter, Auctions: An introduction, *J. Econ. Surv.* **10** (1996), 367–420.