A LEARNING APPROACH TO AUCTIONS

SHLOMIT HON-SNIR*, DOV MONDERER* AND ANER SELA**

* Faculty of Industrial Engineering and Management,
Technion–Israel Institute of Technology, Haifa , Israel
** Department of Economics, Bonn University, Germany

October 1996 (first version: September 1996)

We thank the participants of the microeconomics workshop, held at Northwestern University on August 1996 for their useful comments, and in particular we thank Larry Samuelson for Example 3. We also thank Ezra Einy, Ron Holzman, Doron Sonsino, and Benyamin Shitovitz 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.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

Running Head: Learning in Auctions

Mailing Address: Dov Monderer
Faculty of Industrial Engineering and Management
The Technion
Haifa 32000
Israel

Abstract. 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 for sufficiently large 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, D83, D44.

Contents.

1. Introduction to Learning Theory. Consider a decision maker who is about to play a repeated game in strategic (normal) form. He knows the sets of actions that are available to the players and his own utility function, defined on the set of action profiles, but he does not know the utility functions of the other players. In classical game theory such a situation is modeled as a repeated game with incomplete information. In this model, the decision maker imagines a set S of states with a prior probability μ and an indexed set of one-stage games $(G_s)_{s \in S}$. The description of each G_s includes a list of players, a list of actions for each player, a list of utility functions, and a description of the information which will be available to each player in G_s after each stage. He further imagines a collection of random variables x^i, defined on S, one for each player (including himself) which represent the information channels of the players. He assumes that the above is commonly known to all players, and he proceeds to choose a strategy in the repeated game. But before he does so, he views himself as an outsider and tries to guess the strategy of each participant under the assumption that each other player does the same. This "simple" brain game leads him to a conclusion that there exists only one reasonable strategy for each player (including himself). The vector of these strategies necessarily forms a Nash equilibrium. He finally begins to play using his own selected strategy[1], and his private information which he considers as the realization of his information channel.

In most of the economic literature of the last three decades Nash equilibrium is used as the basic solution concept, and it is therefore implicitly assumed that decision makers actually act as described above[2].

Lately, game theory has suggested learning theory as a new approach for describing the behavior of decision makers[3]. Kalai and Lehrer (1993b) extended Blackwell and Dubins' theorem (1962) to multi-person games. They deal with players who precisely imagine the whole setup described above, but who are not completely rational, as they do not necessarily succeed to correctly guess their opponents' strategies. Each player has a prior belief over the other players' strategies and he updates the beliefs in a Bayesian manner[4]. Under

[1] Note, that a strategy is a function that assigns an action to each possible history.

[2] Even the early literature of bounded rationality (see e.g., Neyman (1985) and Rubinstein (1986)), which was motivated by the feeling that game theory demands too much from human beings, dealt only with the relaxation of the concept of strategy. It took into account that a strategy must be simple (i.e., computerized), but it did not change the above process of choosing the strategy.

[3] Other approach to bounded rationality deals with the relaxation of the "common knowledge " (Aumann (1976)) assumption (see e.g., Rubinstein (1989), Monderer and Samet (1989), Binmore and Brandenburger (1990), Aumann and Brandenburger (1995), and the references there.

[4] See Jordan (1991) for another Bayesian approach to learning in Game Theory.

certain conditions on the prior beliefs, Kalai and Lehrer proved that eventually the players learn the true strategies. Although the agents in Kalai and Lehrer are not completely rational, their logical and computational capabilities are evidently highly rated, and so are those of the players in models of self-confirming equilibrium, subjective equilibrium, and subjective worlds[5].

In contrast, "naive" learning theory deals with agents with modest capabilities. There are two main approaches in this theory . The first approach, dating back to Darwin (1859), and in game theory to von Newmann and Brown (1950), is supported by biological and psychological observations showing that agents tend to use reinforcement strategies. In such a strategy a player's tendency to use a certain action depends positively on the success of this action in the past[6].

The second approach in naive learning theory is the belief-based learning (Fudenberg and Kreps (1993) and Milgrom and Roberts (1991))[7]. In models that use this approach, it is argued that people are myopic and tend to form beliefs about their opponents' next move, based on the history generated by the opponents[8]. The basic belief-based learning model is the fictitious play (Brown (1951)), in which every player believes that the next action profile of his opponents is distributed according to the empirical distribution of their past action profiles. This model has a lot of variations including bounded recall, weighted empirical distributions, and stochastic noise.

In this paper we analyze repeated first-price auctions in the context of belief-based learning.

2. Introduction to the Paper. This work deals with a first-price sealed-bid auction of a single item. Such type of auction, as well as many other types, has been extensively used as selling mechanism, and has been the subject of an intensive theoretical research

[5] See Fudenberg and Levine(1993), Kalai and Lehrer (1993a), Greenberg (1996), and the list of references there .

[6] This motivated the numerous literature on replicated dynamics (See the excellent exposition in Hofbauer and Sigmund (1988) and the list of references there), the case-based decision theory of Gilboa and Schmeidler (1992), and the experiment analysis of Roth and Erev (1995) and Erev and Roth (1996). Erev and Roth analyzed the results of several experiments and concluded that the behavior of players in a repeated game with a unique non trivial mixed strategy equilibrium can be explained by the utilization of a particular reinforcement strategy.

[7] Actually this approach originated in Brown (1951). It seems however that Brown thought of his strategy as a technical device for computing values of 2-person zero-sum games , and not as a strategy that real agents may actually use (as the name "fictitious play" suggests).

[8] Belief-based learning is used in evolutionary economics and game theory, alongside the reinforcement learning (see e.g., Kandori, Mailath, and Rob (1993), Young (1993), and Kandori and Rob (1995)).

in economics and operations research[9]. Much of this research has focused on the equilibrium analysis of the corresponding one-shot Bayesian game[10]. Other research efforts have been devoted to auction design, based on equilibrium analysis[11]. This work differs from previous ones in three main aspects: a) It discusses discrete models , b) It deals with repeated auctions with incomplete information, 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. 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[12]. In our main result (Theorem A in Section 5) we prove under mild assumptions concerning tie-breaking rules, that for sufficiently large time, the players play an equilibrium of the one-shot auction in which players' types are common knowledge. In Section 5 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 are approaching 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 that was communicated to us by Larry Samuelson. Section 7 is devoted to some other learning schemes that seem natural in the context of auctions, and in both Section 7 and Section 8 we provide some remarks and open problems.

3. Repeated (Discrete) First-Price Sealed-Bid Auction. Let $N = \{1, 2, \ldots, n\}$ be the set of players. In the one-shot auction $A(v^1, v^2, \ldots, v^n)$, 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.

---

[9] It is very difficult to list the numerous papers on auctions. The reader is referred to to the surveys of Stark and Rothkopf (1979), Milgrom (1985, 1987), McAfee and McMillan (1987), Wilson (1992), and to the evolving recent literature concerning auctions on spectrum rights (e.g., McMillan (1992) and Cramton (1995).

[10] See e.g., (in addition to the above mentioned surveys) the early work of Vickrey (1961) and the more recent and comprehensive approach to equilibrium of Milgrom and Weber (1982).

[11] See Harris and Raviv (1981), Myerson (1981), and Riley and Samuelson (1981).

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

The action set of each player is the set $Z_+ = \{1, 2, \dots\}$ 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 $\frac{1}{w}(v^i - x^i)$, 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 (though they play a crucial role in the standard equilibrium analysis). In the repeated auction, Nature chooses the types $v^1, v^2 \dots, v^n$, and the auction is repeatedly conducted. The repeated auction is denoted by $RA(v^1, v^2 \dots, v^n)$.

4. 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 $N = \{1, 2, \dots, n\}$. 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[13]. 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 $f^i : \cup_{t=0}^{\infty} H_t \to S^i$. For a finite set X, $\Delta(X)$ denotes the set of probability measures over X. A *belief function* for i is a function $B^i : \cup_{t=e^i}^{\infty} H_t \to \Delta(S^{-i})$, 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 t^{th} joint action of all other players, after he observes the history $h_{t-1} = (x_1, \dots, x_{t-1})$. Player i generates believes only after he observes at least e^i action profiles[14]. Let BR^i be the pure best response correspondence of Player i. 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 that those of Milgrom and Roberts (1991)— For an infinite history $h = (x_1, x_2, \dots)$, we denote $(x_1, \dots, x_t)$ by $h_{[t]}$. A learning scheme (B^i, f^i) is *adaptive* if it satisfies the following three conditions for every infinite history $h = (x_1, x_2, \dots)$:

[13] 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, i can compute his best-response correspondence.

[14] e^i can be interpreted as the length of the experimenting period. It can be used to capture the concept of prior beliefs.

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 $h_s = (z_1, \ldots, z_s) \in H_s$ if $z_t \neq x^{-i}$ for every $1 \leq t \leq s$, 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 $1 \leq s \leq t-1$ 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 $B^i(h_{[t-1]})(x^{-i}) > 0$.

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 an action that has never been used. Condition $AD3$ means that Player i does not ignore recent information. (B^i, f^i) is a *fictitious play (FP) learning scheme* if for every $t > e^i$ and for every history h_{t-1}

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

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

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

If $w_t = 1$ for every $t \geq 1$, then the associated learning scheme is the FP learning scheme. Note that every GFP learning scheme is adaptive, because $\lim_{s \to \infty} \frac{w(T)}{W(T+s)} = \infty$. We say that a learning scheme (B^i, f^i) has a bounded recall if there exists $1 \leq m^i \leq e^i$ 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 $(z_1, \ldots, z_{m^i}) \in H^{m^i}$ and for every $h_{t-m^i-1}, \bar{h}_{t-m^i-1} \in H^{t-m^i-1}$

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

and

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

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 $t - m^i \leq s \leq t-1$ such that $x_s^{-i} = x^{-i}$, where $h_{t-1} = (x_1, \ldots, x_{t-1})$).

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 4.1. *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 $m^i \geq 1$ is the minimal positive integer which satisfies $BR1$ and $BR2$. The following simple lemma gives a useful principle. This principle has already been proved in other versions by Fudenberg and Kreps (1993), and by Monderer and Sela (1993) (who call it the "stability principle").

Lemma 4.2. *Consider a repeated game as described above. Let $h = (x_1, x_2, \dots)$ 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 $t \geq T_0$. 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.

5. Belief-based Learning In Auctions—Main Theorems.

We proceed to analyze the paths generated by players in the repeated auction $RA(v^1, v^2, \dots, v^n)$, 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 $x^i \geq v^i$ 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 $v^i - 1$.

Proposition 1. *Let $RA(v^1, v^2, \dots, v^n)$ 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 $A(v^1, v^2 \dots, v^n)$, such that $x_t = x$ for every $t > T_0$.*

Proof. Denote the recall size of Player i by m^i. Assume without loss of generality that Nature chooses the types in a non-increasing order. That is, $v^1 \geq v^2 \geq \dots \geq v^n$. Let $h = (x_1, x_2, \dots)$ 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 t^{th} actions of the other players. That is, $p_t^j = B^j(x_1, \dots, x_{t-1})$. Let $p_t^j[b]$ denotes the p_t^j-probability that the maximal bid of all other players is b, and let

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

where $|B|$ denotes the number of players in B. 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

$$\frac{1}{n} p_t^j[b] \leq q_t^j[b] \leq \frac{1}{2} p_t^j[b]. \tag{5.1}$$

We need the following claim.

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

Proof of claim 1. If $v^1 = v^2$, then $v^j \leq v^2$ for every player j and therefore the the assertion follows from $TB2$. If $v^1 > v^2$, then all players in $N \setminus \{1\}$ bid less than v^2 by $TB2$. Therefore, for $t > e \geq e^1$, 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 $t > T_0$, $x_t^j \geq v^2 - 2$ and in addition, if M^1 is a singleton, then $y_t(2) \geq v^2 - 2$ for each such t.[15]

[15] 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, then the players may generate the path : $(5, 4, 1), (5, 4, 1), \dots$.

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

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

$$p_t^j[1] + \cdots + p_t^j[k-1] \geq \frac{v^j - k - 2}{2} p_t^j[k], \tag{5.2}$$

where the left-hand side of (5.2) equals zero when $k = 1$.

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

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

As $q_t^j[k+1] \geq 0$ and by (5.1), $q_t^j[k] \leq \frac{1}{2} p_t^j[k]$, the last inequality yields

$$p_t^j[1] + \ldots p_t^j[k-1] \geq ((v^j - k - 1) - \frac{v^j - k}{2}) p_t^j[k].$$

Hence, (5.2) is obtained by manipulating the right-hand side of the previous inequality.

The proof of the next claim is obvious.

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

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

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

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

As $v^1 > 3$, $p_t^j[1] = 0$. This implies by (5.1) that $q_t^j[1] = 0$. As j weakly prefers 1 to $v^1 - 1$, Claim 3 yields:

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

This implies that one of the other players, say player i chose $x_s^i \geq v^j$, for some $t - m^j \leq s \leq t-1$, 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 that

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

Therefore $x_s^1 = v^2$ for every $t - m^i \leq s \leq t-1$. As by $TB2$ $x_{t-1}^i < v^2$, x_{t-1}^i is a best response to Player i' 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 an 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 $TB1$, $x_{t-2}^i = 1$. Continuing recursively, we show that for every $i \in M^2$, Player i bids 1 in the stages s, $t - m^1 - 1 \leq s \leq t-2$. Therefore, at time $t-1$ Player 1 plays v^2 when he observes an 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 $k-1$, $2 \leq k \leq v^2 - 3$, 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, (5.2) holds. By the induction hypothesis the probability that the maximal bid of all other players is less or equal $k-1$ equals zero, hence the left-hand side of (5.2) is zero and therefore

$$0 \geq \frac{v^j - k - 2}{2} p_t^j[k].$$

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

$$0 = (v^j - k)q_t^j[k] \geq p_t^j[1] + \cdots + q_t^j[v^j - 1].$$

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 not a singleton. Let $i_t \in M^2$ be a player with $x_t^{i_t} = y_2(t)$. Since i_t bids k at time t, then for every $i \in M^2$, $x_t^i \leq k$. Let $i \in M^2$, and denote x_t^i by τ. As i weakly prefers τ to $\tau + 1$ we get (as before) that $p_t^i[\tau] = 0$. Since i weakly prefers τ

to v^2-1, we get as before that $p_t^i[1]+\cdots+q_t^i[v^2-1]=0$. Therefore, Player 1 played v^2 in the last m^i moves. By $TB1$, for every $i \in M^2$, $x_{t-1}^i = x_t^i$. 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 $s > T_0$, $x_s^1 \geq v^2-2$ and $y_s(2) \geq v^2-2$. We show that for $t > T_0+e$, $x_t^1 \geq v^2-1$. 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

$$0 \geq \frac{v^1-v^2}{2} p_t^1[v^2-2],$$

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

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

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 $x^i = v^2-1$, and for every player j, $x^j \leq v^j-1$.

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 $x_{T^*+1}^i = v^2-1$ for every $i \in M^2$. Therefor 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 observes 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 $i \in M^2$, and $x^j \leq v^j-1$ for every player j.

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 $j \neq 1$, $x^j = x_{\bar{T}+e}^j$, 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 $t \geq T^* + 1$. Therefore for every $t > T^* + m^1$, Player 1 observes an 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 an history (if $|M^2| = 1$, then $v^2 - 1$ is also a best reply). However, for every $t \geq T^* + m^1 + 1$, $x_t = x$, where $x^i = v^2 - 1$ for every $i \in M^2$, $x^j \leq v^j - 1$ for every j, 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 $x^i \leq v^i - 1$ 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

$$E_1 = p_t^j[v^1 - 2] + q_t^j[v^1 - 1] \geq 2q_t^j[v^1 - 2] + q_t^j[v^1 - 1] > 2q_t^j[v^1 - 2] = E_2.$$

If $p = 1$, then 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, $q_t^j[v^1 - 2] \leq \frac{1}{3}p_t^j[v^1 - 2]$. Therefore $E_2 \leq \frac{2}{3}E_1 < E_1$.

Case 2.2 ($|M^1| = 2$): In this case both players in M^1 bid either $v^1 - 1$ or $v^1 - 2$ for sufficiently large stage. If one of them bids $v^1 - 1$ once, 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 $x^j \leq v^j - 1$ for every player j. It may be, however, that both players play $v^1 - 2$ for ever , 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 thereafter make $v^1 - 1$ a strictly best reply for the players in M^1. □

Proposition 2. *Let $RA(v^1, v^2, \ldots, v^n)$ 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 $A(v^1, v^2 \ldots, v^n)$, such that $x_t = x$ for every $t > T_0$.*

Proof. Assume without loss of generality that Nature chooses the types in a non-increasing order. That is, $v^1 \geq v^2 \geq ... \geq v^n$. Let $h = (x_1, x_2, \dots)$ 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 $\{1, 2, \dots, v^2 - 3\}$ only finitely many times. Moreover, if M^1 is a singleton and*

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

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

Proof of Claim 4. This claim is obvious if $v^2 < 4$, thus we proceed to prove it assuming that $v^2 \geq 4$. We prove by induction on $1 \leq k \leq v^2 - 3$, that every $j \in M^1$ makes a bid in $\{1, \dots, k\}$ 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 $\{1, \dots, k\}$ 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 $w \geq v^1$, $p_t^j[w] > 0$. Therefore, by $AD2$, there exists a player i such that for some $1 \leq s \leq t - 1$, $x_s^i = w \geq v^j$, 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 $e < t_1 < t_2 < \dots$. 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 $1 \leq s \leq t_l - 1$. As $\lim_{l \to \infty} t_l = \infty$, $x_s^1 = v^2$ for every $s \geq 1$, contradicting $A1$.

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

$$w_j(t_l - 1)(p^j_{t_l}[1] + \ldots p^j_{t_l}[k-1]) \geq w_j(t_l - 1)(\frac{v^1 - k - 2}{2} p^j_{t_l}[k]).$$

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:

$$w_j(t_l - 1)(v^1 - k)(p^j_{t_l}[1] + \cdots + q^j_{t_l}[k]) \geq w_j(t_l - 1)(p^j_{t_l}[1] + \cdots + q^j_{t_l}[v^1 - 1]).$$

Since the left-hand side of this inequality is bounded, so is the right-hand side. This implies that the integers $\{1, \ldots, v^1 - 1\}$ 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 $e < t_1 < t_2 < \ldots$. At time t_l Player i weakly prefers k to $k+1$, thus we conclude, as in the first part of the k^{th} step, that there exists M such that

$$w_i(t_l - 1)p^i_{t_l}[k] \leq M \quad \text{for every } l \geq 1. \tag{5.3}$$

Since at stage t_l Player i weakly prefers k to $v^2 - 1$, Claim 3 and (5.3) yield for every $l \geq 1$

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

The last inequality implies that there exists T_0 such that Player 1 bids v^2 for every $t \geq T_0$ 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 $e < t_1 < t_2 \ldots$. 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 $l \geq 1$

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

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:

$$w_1(t_l-1)(v^1-v^2+2)(p^1_{t_l}[1]+\cdots+q^1_{t_l}[v^2_2]\geq w_1(t_l-1)(p^1_{t_l}[1]+\cdots+q^1_{t_l}[v^1-1]).$$

Since the left-hand side of this inequality is bounded, so is the right-hand side. This implies that the integers $\{1,\ldots,v^1-1\}$ 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 $\bar{T}$ such that for every $t\geq\bar{T}$, $x_t=x$, where x is in equilibrium in the auction $A(v^1,v^2,\ldots,v^n)$ satisfying $x^1=v^2$, there exists $i\in M^2$ with $x^i=v^2-1$, and $x^j\leq v^j-1$ for every player j.

Assume $A1$ is not satisfied. Then there exists $T_0>e$ such that Player 1 bids v^2 for $t\geq T_0$. Therefore, by $TB1$, for each player $j\neq 1$, $x^j_t=x^j_{T_0}$ for every $t\geq T_0$. If $x^i_{T_0}<v^2-1$ 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 $x^j\leq v^j-1$ 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 $t\geq T_0$, 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 that 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 $x^j\leq v^j-1$ for every Player j.

Case 2 ($|M^1| \geq 2$)**:**

$v^1 = v^2$. 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 $|M^1| \geq 2$. 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 $T^* < t_1 < t_2 < \ldots$. Then for every $l \geq 1$, $E^j_{t_l}(v^1 - 2) \geq E^j_{t_l}(v^1 - 1)$. Hence,

$$2(p^j_t[1] + \cdots + q^j_t[v^1 - 2]) \geq p^j_t[1] + \cdots + p^j_t[v^1 - 2] + q^j_t[v^1 - 1].$$

Therefore,

$$(p^j_{t_l}[1] + \cdots + p^j_{t_l}[v^1 - 3]) \geq p^j_{t_l}[v^1 - 2] - 2q^j_{t_l}[v^1 - 2] + q^j_{t_l}[v^1 - 1]. \tag{5.4}$$

By (5.1), $p^j_{t_l}[v^1 - 2] \geq 2q^j_{t_l}[v^1 - 2]$. Therefore,

$$(p^j_{t_l}[1] + \cdots + p^j_{t_l}[v^1 - 3]) \geq +q^j_{t_l}[v^1 - 1]. \tag{5.5}$$

Multiplying both sides of (5.5) by $w^j(t_l)$ gives a bounded left-hand side (when l varies). Therefore a bounded right hand-side. Thus there exists T_0 such that $x^d_{t_l} = v^2 - 2$ 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 $x^i \leq v^i - 1$ 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 2 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 $x^i \leq v^i-1$ for every player i, or $x^1=x^2=v^1-2$ and $x^i \leq v^i-1$ for every player i. □

Combining the ideas in the proofs of Proposition 1 and Proposition 2, we can derive the proof of the following combined theorem.

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 $A(v^1, v^2 \ldots, v^n)$, such that $x_t = x$ for every $t > T_0$.*

We end this section with a remark concerning a possible seemingly shorter proof of Proposition 2. A path $(x_1, x_2, \ldots)$, in S is a *better-reply* path if for every $t \geq 1$ 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 (1993) proved that if all players use a JFP learning scheme and 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 JFP 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[16].

Example 1.

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

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

6. 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 $TB1$.

[16]Monderer and Sela 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 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

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

be the sequence of 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[17]. Note, however, that the players in Example 2 may generate a non-converging belief sequence. For example they may generate the path:

$$(5,4),(5,4),(5,x_3^2),(5,4),(5,4),(5,x_6^2),\ldots, \tag{6.1}$$

where x_{3k}^2 is an arbitrary integer in $\{1,2\}$. Though the belief sequence generated by the path in (6.1) does not necessarily converge, it approaches equilibrium in the sense of Monderer and Shapley (1996): 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 $T \geq 1$, such that (p_t, q_t) is an ε-equilibrium for every $t \geq T$.[18]

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

Proof. Denote the belief sequence by $((p_t, q_t))_{t\geq 2}$. We use the following characterization for approaching to equilibrium given in Monderer and Shapley (1996): $((p_t, q_t))_{t\geq 1}$ is

[17] 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.

[18] Equivalently, for every $\varepsilon > 0$ there exists $T \geq 1$, such that for every $t \geq T$, 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 (1951) (zero-sum games), and Miyasawa (1961) (2×2 games)) prove that the belief sequence is approaching equilibrium and not necessarily converging to equilibrium.

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 $t \geq T_0$. Therefore p is the probability measure concentrated on v^2. As for every $t \geq T_0$, 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 $\{1, \ldots, v^2 - 1\}$ 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 $t \geq T_0$. Since $A1$ holds, Player 1 bids $v^2 - 1$ infinitely many times. Therefore, for sufficiently large t, Player 2's conditional probability of the maximal bid of Player 1 being $v^2 - 1$, given that this maximal 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 larger stage at which Player 1 switches to v^2 and stays with this *bid*, in contradiction to $A1$. Therefor 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 enforces

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 action 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 large 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 player bids $v^2 - 2$ infinitely many times, then by Claim 6, (p, q) is the equilibrium $(\delta_{v^2-2}, \delta_{v^2-2})$. □

In the next example we assume that both players use a FP learning scheme with bounded recall. It is shown that the path of actions that is generated by the players does not stabilize and the belief sequence does not approach equilibrium.

Example 3(Samuelson).

There are 2 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 $n \geq 3$ 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 to 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 $v^1 = 9$, $v^2 = v^3 = v^4 = 5$. assume the players use FP learning schemes. They may generate the following path.

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

The individual empirical distribution vector is converging to (p, q^2, q^3, q^4), where $p(5) = 1$, and for $2 \leq i \leq 4$, $q^i(4) = \frac{1}{4}$ and $q^i(1) = \frac{3}{4}$. It is easily verified that for Player 1, the bid 5 is not a best-response to (q^2, q^3, q^4). Therefore (p, q^2, q^3, q^4) is not in equilibrium.

7. Other Learning Schemes. In this section we discuss belief-based learning schemes, which seem natural in the context of repeated auction, but 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 (1996). 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.[19] 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 (1996) proved that when all players use a AMB learning scheme, and apply the tie breaking rules $TB1$ and $TB2$, then the generated path of action profiles stabilizes on equilibrium.

8. Additional Remarks: Future Research.

Domination: Assume a repeated game which is solvable by successive elimination of strongly 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. One can deduce from Milgrom and Roberts (1991), 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 not known whether a version of this result holds for weakly dominated strategies. Since Hon-Snir (1996) proved that the first-price auction discussed here is solvable by successive elimination of *all* weakly dominated strategies, our result shows that this is indeed the case for our particular model.

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

[19] The belief function in this scheme takes deterministic values. Thorlund-Peterson(1991) discusses such learning scheme applied to the Cournot game.

based learning schemes with such imperfect monitoring[20] will contribute to auction theory.

Optimal Auctions: A study of learning in auctions different from the simple first-price auction analyzed in this paper, may open the gate for a theory of optimal auctions, under the learning assumption rather then the equilibrium assumption.

Reinforcement Learning: In the theory of reinforcement learning, players do not form beliefs about the other players' next move. They are assumed to use 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 stochastic process in S. Hon-Snir (1996) 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 $\{1, 2, \ldots, \bar{V}\}$. She shows that if the number of possible types for each player does not exceed seven (i.e., $\bar{V} \leq 7$), then with probability one, for sufficiently large 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 , though 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. Since this is a very natural assumption we do not actually prove this conjecture.

9. References.

R. J. Aumann, Agreeing to Disagree, *Annals of Statistics* **4** (1976), 1236-1239.

R. J. Aumann and A. Brandenburger, Epistemic Conditions For Nash Equilibrium, *Econometrica* **55** (1995),1161-1180.

K. Binmore and A. Brandenburger, Common Knowledge and Game Theory, in K. Binmore (Ed.), *Essays on the Foundations of Game Theory* (1990), Basil Blackwell, Oxford.

[20] Like the AMB learning scheme described in Section 7.

D. Blackwell and L. Dubins, Merging of Opinions With Increasing Information, *Annals of Mathematical Statistics* **38** (1962), 882-886.

G. W. Brown and J. von Neumann, Solution of Games by Differential Equations, in: Contributions to the Theory of Games, *Annals Mathematics Studies* (Eds., H.W. Kuhn and A. W. Tucker) **24** (1950), 73-79.

G. W. Brown, Iterative Solution of Games by Fictitious Play, *in* "Activity Analysis of Production and Allocation", Wiley, New York, 1951.

P. C. Cramton, Money Out of Thin Air: The Nationwide Narrowband PCS Auction, mimeo, 1995. *Journal of Economics and Management Strategy*, forthcoming.

Ch. Darwin, On the Origin of Species by Means of Natural Selection, (1859), Cambridge London. Reprinted in Harvard University Press (1964).

I. Erev and A.E. Roth (1996).

D. Fudenberg and D. Kreps, Learning Mixed Equilibria, *Games and Economic Behavior* **5** (1993), 320-367.

D. Fudenberg and D. Levine, self-Confirming Equilibrium, *Econometrica* **61** (1993), 523-545.

J. Greenberg, Towering Over Babel: Worlds Apart–But Acting Together, McGill University (1996).

M. Harris and A. Raviv, Allocation Mechanisms and the Design of Auctions, *Econometrica* **49** (1981),1477-1499.

J. Hofbauer and K. Sigmund, The Theory of Evolution and Dynamical Systems, Cambridge University Press (1988), New York.

S. Hon-Snir, (1996), M.Sc. Thesis and Ph.D. Thesis, The Technion.

J. S. Jordan, Bayesian Learning In Normal Form Games, *Games and Economic Behavior* **3** (1991), 60-81.

E. Kalai and E. Lehrer, Subjective Equilibrium in Repeated Games, *Econometrica* **61** (1993a), 1231-1240.

E. Kalai and E. Lehrer, Rational Learning Leads to Nash equilibrium, *Econometrica* **61** (1993b), 1019-1045.

M. Kandori, G. Mailath, and R. Rob, Learning, Mutation and Long run equilibria in Games, *Econometrica* **61** (1993), 27-56.

M. Kandori and R. Rob, Evolution of Equilibria in the Long Run: A general Theory and Applications, *Journal of Economic Theory* **65** (1995),383-414.

R. P. McAfee and J. McMillan, Auctions and Bidding, *Journal of Economic Literature* (1987), 699-738.

J. McMillan, Selling Spectrum Rights, *Journal of Economic Perspective* **8** (1992), 145-162.

P. R. Milgrom, The Economics of Competitive Bidding: A Selective Survey, in L. Horwicz, D. Schmeidler and H. sonnenschein, eds., *Social Goals and Social Organizations* (1985), Cambridge University Press, Cambridge.

P. R. Milgrom, Auction Theory, in : T. Bewley, ed., *Advances in Economic Theory: Fifth World Congress* (1987), Cambridge University Press, Cambridge.

P. R. Milgrom and R. J. Weber, A Theory of Auctions and Competitive Bidding, *Econometrica* **50** (1982),1089-1122.

P.R. Milgrom and J. D. Roberts, Adaptive and Sophisticated Learning in Normal Form Games, *Games and Economic Behavior* **3** (1991), 82-100.

K. Miyasawa, On the Convergence of the Learning Process in a 2×2 Non-Zero-sum Two Person Game, *Economic Research Program, Princeton University, Research Memorandum No. 33* (1961),

D. Monderer and D. Samet, Approximating Common Knowledge with Common Beliefs, *Games and Economic Behavior* **1** (1989), 170-190.

D. Monderer and A. Sela, Fictitious Play and No-Cycling Conditions, mimeo, 1993.

D. Monderer and L. S. Shapley, Fictitious Play Property for Games with Identical Interests, *Journal of Economic Theory* **1** (1996), 258-265.

R. B. Myerson, Optimal Auction Design, *Mathematics of Operations research* **6** (1981), 58-63.

A. Neyman, Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoner's Dilemma, *Economics letters* **19** (1985), 227-229.

J. G. Riley and W. F. Samuelson, Optimal Auctions, *American Economic Review* **71** (1981), 381-392.

J. Robinson, An Iterative Method of Solving a Game, *Annals of Mathematics* **54** (1951), 296-301.

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

A. Rubinstein, Finite Automata Play the Repeated prisoner's dilemma, *Journal of eco-*

nomic Theory **39** (1986),83-96.

A. Rubinstein, The electronic mail Game: strategic behavior under “almost common knowledge”, *American Economic Review* **79** (1989), 385-391.

L. S. Shapley, Some Topics in Two-Person Games, *in* “Advances in Game Theory” (M. Dresher, L.S. Shapley, and A.W. Tucker, Eds.) 1-29, Univ. Press, Princeton, NJ, 1964.

R. M. Stark and M. H. Rothkopf, Competitive Bidding: A Comprehensive Bibliography, *Operations Research* **27** (1979), 364-390.

L. Thorlund-Petersen, Iterative Computation of Cournot Equilibrium, *Games and Economic Behavior* **2** (1990), 61-95.

W. Vickrey, Counterspeculation, Auctions, and Sealed Tenders, *Journal of Finance* **16** (1961), 8-37.

R. Wilson, Strategic Analysis of Auctions, in: R. J. Aumann and S. Hart, eds., *Handbook of Game Theory, Volume 1* (1992), Elsevier Science Publishers.

P.Young, The Evolution of Conventions, *Econometrica* **61** (1993), 57-84.