Ex-post Regret Heuristics under Private Values (I): Fixed and Random Matching

Rene Saran† Roberto Serrano‡

This version: May 2014

Abstract

In contexts in which players have no priors, we analyze a heuristic process based on ex-post regret as a guide to understand how to play games of incomplete information under private values. The conclusions depend on whether players interact within a fixed set (fixed matching) or they are randomly matched to play the game (random matching). The relevant long run predictions are minimal sets that are “closed under same or better reply” operations. Under additional assumptions in each case, the predictions boil down to Nash equilibria, ex-post equilibria or minimax regret equilibria. These three paradigms exhibit nice robustness properties in the sense that they are independent of beliefs about the exogenous uncertainty of type spaces. The results are illustrated in several applications, including second-price auctions, first-price auctions and Bertrand duopolies.

Keywords: Fixed and Random Matching; Incomplete Information; Ex-Post Regret Heuristics; Nash Equilibrium; Ex-Post Equilibrium; Minimax Regret Equilibrium; Second-Price Auction; First-Price Auction; Bertrand Duopoly.

JEL: C72; C73; D82; D83.

*This paper supersedes a previous draft (Saran and Serrano, 2010). We thank the editor, two anonymous referees, Larry Ausubel, Antonio Cabrales, Peter Cramton, Emel Filiz-Ozbay, Sergiu Hart, Larry Samuelson, Bill Sandholm, Peyton Young and seminar audience at Rutgers, Maastricht, Ben-Gurion, Maryland, Kalai Conference (Jerusalem), ASSET 2007 (Padua), NASMES 2008 (Pittsburgh) and Game Theory World Congress (Evanston) for useful comments and suggestions. A special acknowledgment goes to Ronald Peeters for his help with the simulations. Serrano thanks CEMFI in Madrid for its hospitality.

†Social Sciences Division, Yale-NUS College, 6 College Avenue East #07-19, Singapore 138614. Email address: rene.saran@yale-nus.edu.sg; Tel: +65 6601 3380; Fax: +65 6773 6932
‡Department of Economics, Brown University, Providence, RI 02912, U.S.A. Email address: roberto_serrano@brown.edu
1 Introduction

The standard analysis of games with incomplete information relies on the notion of Bayesian equilibrium, and has led to important theoretical constructions. Moreover, equilibrium insights can sometimes be applied with success to real-world problems. However, the use of equilibrium also has important limitations, as it assumes the existence of a common-knowledge type space to describe the underlying uncertainty. When for instance one talks to game theorists that have provided advice on how to bid in real auctions, this is often a major stumbling block: we may calculate an equilibrium of the given auction, but in the absence of common-knowledge of strategies, type spaces and prior beliefs, it is implausible to expect the equilibrium to be played. How should one bid then? More generally, how does a player behave in a game of incomplete information? Our analysis in the current paper provides several possible answers, which should be treated as a guide to behavior in our benchmarks.

In agreement with recent trends in the robust analysis of game theory with incomplete information, we shall deemphasize the role of beliefs and turn to ex-post considerations. In fact, we shall propose an alternative paradigm to equilibrium theory by going all the way to the other extreme: without specifying any prior distribution, we shall analyze a heuristic model of behavior based on ex-post regrets.

We apply our ex-post regret heuristic process to games in which each player knows his or her own payoff function. This makes our results applicable to private-values incomplete-information games. We assume large (but finite) multiple populations – one for each type of each player –, from which players are drawn to play the game. The analysis will be performed both under fixed matching and random matching of the players.

Under fixed matching, a player is selected from each population and that set of players stays matched to play the game infinitely often. In the simplest version of the ex-post regret heuristic rule, each player starts by choosing his or her first action arbitrarily. After that first period, his or her last action is repeated with positive probability (inertia), but the player also switches to other actions with positive probability if and only if he or she regrets not having used them in the last period. That is, the player compares the payoff obtained in the last period and wonders what would have happened, ceteris paribus, had he or she used a different action.\footnote{See Saran and Serrano (2012) for an extension of the analysis to regret matching with respect to any finite memory, as opposed to only one-period memory, studied here.} The player feels regret for not having used those actions that would have strictly
increased his or her last period’s payoff, and switches to playing them with arbitrary positive probability. This means that, at any given period, a player’s action is either unchanged or modified into one that constitutes a strictly better reply to the last action profile. The ex-post regret heuristic rule generates a finite Markov process on the set of action profiles that can be played by the matched players.

In general finite strategic games, we show that the recurrent classes of the Markov process under fixed matching correspond to minimal sets of action profiles that are “closed under same or better replies” (minimal CUSOBR sets). We illustrate this result with the fashion game. Minimal CUSOBR sets are related to the minimal product sets that are “closed under weakly better replies” of Ritzberger and Weibull (1995). Any of the Ritzberger-Weibull closed sets contains one minimal CUSOBR set, sometimes properly, since the “(weakly) better-reply” correspondence contains the “same or (strictly) better-reply” correspondence. However, we find that some minimal CUSOBR sets are not necessarily product sets, and sometimes they need not be included in any closed set à la Ritzberger and Weibull.2

For weakly acyclic games under better replies, the result is stronger. All the recurrent classes of the Markov process under fixed matching are singleton, and the set of absorbing states coincides with the set of pure Nash equilibria of the complete-information game corresponding to the types of the players that have been matched. Thus, in spite of the initial incomplete information, the set of players, through the experience developed by repeatedly playing the game together, arrive at Nash equilibrium play. The process converges with probability 1 to one of the pure Nash equilibria; which of them is selected will depend on the initial condition. Throughout the paper we use several running examples to illustrate our results, including the second-price auction, the first-price auction and Bertrand duopoly competition. We show that these three are weakly acyclic games under better replies.

Under random matching, an independent draw from each player’s population takes place every period. The selected players play for one period and then return to the pool. We conceive a game of incomplete information associated with the random matching model. A strategy for a player in this game is a function specifying for each type of that player, the distribution of actions amongst the corresponding (finite) population. For a given player, two strategies are adjacent if exactly one type of that player has switched his or her action, while all other types have not, in the two strategies in question. Now the ex-post heuristic rule generates a finite Markov

---

2Sets that are closed under weakly better replies are curb (closed under rational behavior) sets (Basu and Weibull, 1991). CUSOBR sets are not, as some better replies are excluded.
process on the set of strategy profiles of the incomplete information game.

The recurrent classes of the Markov process under random matching coincide with those minimal sets of strategy profiles of the incomplete information game that are “closed under same or adjacent ex-post better replies” (minimal CUSOAEBR sets). For any strategy profile, the “same or adjacent ex-post better reply” correspondence consists of the same strategy profile – in which no player has changed his or her action – and those adjacent strategy profiles that are obtained when in some ex-post event, some types of the players switch to (strictly) better replies. We show that singleton CUSOAEBR sets are equivalent to pure ex-post equilibria of the incomplete information game.

For the second-price auction, we show that all minimal CUSOAEBR sets are singleton, and therefore each recurrent class of the Markov process under random matching is an absorbing state that corresponds to a pure ex-post equilibrium of the incomplete information game. However, in general games, minimal CUSOAEBR sets can be rather large. We illustrate this by exhibiting the unique minimal CUSOAEBR sets for the first-price auction and Bertrand duopoly game. Almost every strategy profile is part of such a minimal set, which is therefore uninformative in terms of an economic prediction. In an attempt to gain determinacy, we turn to stochastic monotone dynamics in which the probability of switching to an action is an increasing function of the associated regret and in which the players are allowed to “make mistakes” with small probability. Taking limits as mistakes probability vanishes would not help, however, since the problem is the large recurrent class of the unperturbed process, as just described.

Instead, we propose an approach in which taking another limit is possible and yields an interesting answer. We take the limit as the switch probabilities vanish.\(^3\) If one performs such a limiting exercise, one obtains a selection of the recurrent classes of the associated unperturbed process, which happens to be a snapshot of the game in which the behavior of all agents is fixed, and hence the set of recurrent classes simply coincides with the set of states of the process (equivalently, the set of strategy profiles of the incomplete information game). Thus, the exercise performed in this part of the paper is the long-run prediction in a random matching context in which the adaptive rule is the perturbed monotone ex-post regret heuristics, but where switches are unlikely events. Such an application of stochastic stability provides an extremely

\(^3\)In our companion paper (Saran and Serrano, 2014), we perform a rather comprehensive analysis of the ex-post regret based stochastic monotone dynamics applied to $2 \times 2$ games, a class that has been traditionally studied using other learning approaches. The stochastic stability results in both papers are robust to small nonnegligible probabilities, as confirmed via simulations.
powerful refinement in some cases. For the first-price auction and Bertrand duopoly, a unique strategy profile is selected. In the first-price auction, each player bids half his valuation. In Bertrand duopoly, each firm uses that price that, if it were the only firm in the market, would yield half the monopoly profit. In both cases, these actions are such that the maximum gain from increasing the bid/price equates the maximum gain from decreasing it. It follows that the strategy profiles selected by stochastic stability are minimax regret equilibria (Saran and Serrano, 2014; also see Hyafil and Boutilier, 2004). Our heuristic foundation of minimax regret equilibria in the first-price auction and Bertrand duopoly game is a consequence of the inverse of the regrets being the exponents of the switch probabilities, which go to zero in the stochastic stability analysis. We view this as a valid economic prediction in the absence of information about the exact distribution of types in the population. Savage (1951) was the first to propose minimizing the maximum ex-post regret – minimax regret – as a decision-making criterion in problems in which the agent has no specific prior about the states of the world.

Thus, for each version of the model, our study yields the relevant sets that are “closed under same or better replies” as long-run predictions. Under stronger assumptions, play ends up at pure Nash equilibria of the corresponding complete information game (fixed matching in weakly acyclic games), pure ex-post equilibria (random matching in games with only singleton minimal CUSOAE sets, e.g., second-price auction) and minimax regret equilibria (random matching with unlikely switches and mistakes). With respect to the exogenous uncertainty stemming from type spaces, these three paradigms are belief-independent and provide answers to how the game could be played, which ought to be compared to its Bayesian equilibria.

4We also introduce a test to check for stochastic stability, of interest in its own right, based on the construction of certain weighted cycles. This is especially useful in systems like ours, where the construction of minimal cost rooted trees would require to know the transitions between any pair of states.

5In contrast, Louge and Riedel (2012) show that with two players and uniform distribution of valuations, the Bayesian-Nash equilibrium in which each player bids half of his valuation is not asymptotically stable under payoff monotonic dynamics.

6We also prove this result for \( N \times 2 \) games with dominant actions, generalizing the \( 2 \times 2 \) analysis of Saran and Serrano (2014). At a minimax regret equilibrium, each player uses a strategy that minimizes his or her maximum ex-post regret. Therefore, pure ex-post equilibria are always minimax regret equilibria (because players have no regrets), but the latter set is nonempty whenever there is “sufficient” diversity of types. As discussed in Saran and Serrano (2014), our concept of minimax regret equilibrium differs from that of Hyafil and Boutilier (2004).

7For the reasons outlined above, more research on dynamics and learning/heuristic processes applied to Bayesian games would be desirable. See Dekel, Fudenberg and Levine (2004) and Ely and Sandholm (2005) for early contributions.
1.1 Related Literature

Our processes are related to the no-regret learning literature (see Fudenberg and Levine (1998), Hart (2005) and Sandholm (2009) for different surveys of the area; see also Saran and Serrano (2012) and the references therein). However, our emphasis in the current paper concerns games of incomplete information. As just stated, Savage (1951) was the seminal contribution suggesting minimax regret as a decision rule in contexts in which agents have no prior beliefs. Linhart and Radner (1989) argue that Bayesian-Nash equilibrium does not offer any guidance to behavior when priors are not common knowledge, and therefore study minimax-regret strategies in \(1/2\)-double auctions.

Especially related to our fixed-matching model is Hon-Snir, Monderer and Sela (1998), who model learning in an auction context differently. The players’ valuations are determined first. Then, this fixed set of players repeatedly plays a first-price auction. Players are informed about the profile of bids at the end of each auction. The paper looks at two learning rules: generalized fictitious play and adaptive learning with bounded recall. The main result is that, if all players use either learning rule, then per-period play converges to Nash equilibrium of the one-shot auction in which players’ types are common knowledge.\(^8\)

Our last results use stochastic stability. Most applications of this methodology in noncooperative game theory have been confined to games with complete information.\(^9\) An exception is Jensen, Sloth and Whitta-Jacobsen (2005), which extends the perturbed best-response model in Young (1993) to finite two-player Bayesian games. Compared to our analysis, their players possess much more information, both ex-ante and ex-post.\(^10\)

\(^8\)The particular equilibrium is the one in which the player with the highest valuation wins and pays a price equal to the second-highest valuation. However, this happens because of their assumption that no player bids more than her valuation. We will also get a similar result with this additional assumption.

\(^9\)See Young (1998) for an account of different applications.

\(^10\)They make three assumptions: First, the players know the true distribution of types in the population; second, the types of the matched players are truthfully revealed to everyone at the end of the interaction; and third, for each type of each player, there is a record of the action taken by that type during some past periods in which that type was selected. Under these assumptions, the unperturbed best-response dynamics, appropriately redefined, converges with probability one to some convention, which is a state that is “equivalent” to a strict Bayesian equilibrium of the game – if the latter exists. The perturbations then select among the different strict Bayesian equilibria.
1.2 Plan of the Paper

Section 2 describes the ex-post regret heuristic rule when players make no mistakes and the matching protocols. Section 3 studies the long run behavior under fixed matching, and Section 4 under random matching. Section 5 presents the ex-post regret heuristic rule when players make mistakes and studies the resulting long run behavior under random matching. All of these three sections contain a subsection on applications. Section 6 concludes, and Section 7 collects the proofs of major results. We provide definitions and results related to the concept of stochastic stability in the Appendix.

2 Heuristic Rule with No Mistakes and Matching Protocols

Let $G$ be an $N$-person game of incomplete information with $T_i$ being the finite set of types and $A_i$ being the finite set of actions $a_i$ for each player who plays in position $i \in N$ ($N$ will denote both the set and number of positions). Let $a_{-i}$ denote a profile of actions of all players other than $i$, and $A_{-i}$ be the set of such action profiles. We assume private values, i.e., player $i$’s ex-post payoff $\pi_i(t_i, a_i, a_{-i})$ is a function of only her type $t_i$ and the action profile $(a_i, a_{-i})$. A player’s type is her private information, it is drawn once and for all, and remains constant over time. Note that a player’s type here involves no beliefs regarding other players’ types.

There is a large but finite population of players $P_i$ for each position $i \in N$ (we will also use $i$ to denote a player who plays in position $i$). For each position $i \in N$ and each type $t_i \in T_i$, let $P_{t_i} \subseteq P_i$ be the nonempty subset of players whose types equal $t_i$.

Let $\nu_i$ be any probability distribution over $P_i$ such that each player in population $P_i$ has a positive probability of being selected under $\nu_i$. Players’ interactions will take place under two distinct matching assumptions: fixed matching and random matching:

- **Fixed Matching.** In the initial period, a single player is selected for each position $i$ from population $P_i$ using $\nu_i$. The selected players are matched once and for all, and they repeatedly play the game every period.

- **Random Matching.** In each period, a new player is selected for each position $i$ from population $P_i$ using $\nu_i$. The selected players are matched to play the
In every period, a player in population $P_i$ is identified by her type $t_i$ and her unique action $a_i$ that she plays in the event she is matched. Players’ actions in the first round are chosen arbitrarily.$^{11}$ We shall assume that players adapt their actions from one period to the next using the following ex-post regret heuristic rule: suppose in period $l$ a player with type $t_i$ is choosing action $a_i$.

- If this player is not matched in period $l$, then she does not change her action in period $l + 1$.
- If this player is matched in period $l$ with other players who play $a_{-i}$, then she obtains a payoff of $\pi_i(t_i, a_i, a_{-i})$ in period $l$ game. Pick any $a_i' \in A_i \setminus \{a_i\}$. Had she chosen $a_i'$ instead of $a_i$ in that game, ceteris paribus, her payoff would have been $\pi_i(t_i, a_i', a_{-i})$.$^{12}$

Define $\Delta_i(t_i, a_i, a_i', a_{-i}) = \pi_i(t_i, a_i', a_{-i}) - \pi_i(t_i, a_i, a_{-i})$.

If $\Delta_i(t_i, a_i, a_i', a_{-i}) > 0$, then we refer to this number as type $t_i$’s ex-post regret from using $a_i$ instead of $a_i'$ against $a_{-i}$. In this case, letting $q(\cdot) \in [0, 1]$ as a basic switch probability function, we assume that the player changes her action to $a_i'$ in period $l + 1$ with probability $q(\Delta_i(t_i, a_i, a_i', a_{-i})) > 0$ if and only if $\Delta_i(t_i, a_i, a_i', a_{-i}) > 0$, and she does not switch her action with a positive probability, i.e., $\sum_{a_i' \in A_i \setminus \{a_i\}} q(\Delta_i(t_i, a_i, a_i', a_{-i})) < 1$. Several different specifications

$^{11}$The initial choice of actions determines the recurrent class to which per-period play converges in the long run. However, the initial condition does not matter for the perturbed ex-post regret heuristic rule in Section 5 since stochastically stable states are independent of the initial choice of actions.

$^{12}$In some of our applications, viz., second-price auction, first-price auction and Bertrand duopoly, there is a slight difference between the definitions of payoffs in the game and payoffs in the ex-post regret heuristic rule. Whenever multiple players are tied at the highest bid/lowest price in these applications, we make the standard assumption that the players’ payoffs in the game are the expected payoffs, where the expectation is with respect to the tie-breaking lottery. In contrast, ex-post regrets are calculated after the players have observed the consequence of their actions, i.e., once they have realized their payoffs. The realized payoff is however a random variable because it depends on the tie-breaking lottery in the event that there is tie at the highest bid/lowest price. Therefore, while defining our ex-post regret heuristic rule in these cases, we let $\pi_i(t_i, a_i, a_{-i})$ be one of the possible realized payoffs – instead of letting $\pi_i(t_i, a_i, a_{-i})$ be the expected payoff as we do in the definition of the game. However, there is no discrepancy in the definition of $\pi_i(t_i, a_i', a_{-i})$, i.e., the payoff obtained by the player after she unilaterally switches from $a_i$ to $a_i'$; in the game as well as in the ex-post regret heuristic rule, if after this switch the bid/price $a_i'$ is tied at the highest bid/lowest price, then $\pi_i(t_i, a_i', a_{-i})$ is the expected payoff defined by the tie-breaking lottery.
of the function \( q(\cdot) \) fulfill these properties;\(^\text{13}\) in our analysis, we of course fix \( q(\cdot) \) to be one such function.

All these events entailing switches in actions are independent across players and time.

This heuristic rule satisfies three properties: (i) inertia, i.e., a matched player does not switch her action with a positive probability, (ii) a matched player switches her action in the next period with a positive probability if and only if her ex-post regret from not using that action is positive (in Section 5, we will specify a perturbed version of this heuristic rule in which players make mistakes, i.e., switch their actions even though they do not experience positive ex-post regret), and (iii) this heuristics is “belief free” because switching probabilities are a function of players’ ex-post regrets that do not require players to formulate beliefs over their opponents’ types and actions.

Our concern in the next two sections is the identification of the long-run behavior under fixed and random matching of the players who follow our ex-post regret heuristic rule. The dynamic process under fixed matching \([M_{FM/G}(q)]\) is defined over states that are profiles of actions taken by the fixed set of players. The dynamic process under random matching \([M_{RM/G}(q)]\) is defined over states that specify the distributions of actions in the different populations of players. Both are finite Markov processes, and they are generally not irreducible, but they are still aperiodic, and convergence always obtains to one of their recurrent classes.\(^\text{14}\) We shall therefore state our results in terms of the recurrent classes of the dynamics \( M_{FM/G}(q) \) and \( M_{RM/G}(q) \). When switches to actions happen after an agent has positive ex-post regret, the analysis will reveal which actions are taken in the long run by a fixed set of players (fixed matching) and which distributions of actions are more likely to emerge in the long run in the large populations (random matching).

3 Long-Run Behavior under Fixed Matching

Let \( G(t_1, \ldots, t_N) \) denote the complete information game in which the types of the players are \( \{t_1, \ldots, t_N\} \), the sets of actions are \( (A_i)_{i \in N} \) and the payoffs are given by

---

\(^{13}\)For instance, let \( \Delta^* \) be the maximum ex-post regret across all types of all players and \( A^* \) be the maximum number of actions that any player has. Then any \( q(\cdot) \) such that \( q(x) \in [0, \frac{1}{\Delta^* A^*}] \) and \( q(x) > 0 \iff x > 0 \) will fulfill these properties. We provide another set of specifications in our stochastic stability analysis.

\(^{14}\)See the Appendix for basic concepts related to finite Markov processes.
the functions \((\pi_i(t_i, \cdot, \cdot))_{i \in N}\). This will be the underlying game once a fixed matching happens, though recall that our players do not know the other players’ types.

Let \(A = \prod_{i \in N} A_i\). For any \((a_i, a_{-i}) \in A\), the set of same-or-better replies for the player of type \(t_i\) is

\[
R_{t_i}(a_i, a_{-i}) = \{a'_i \in A_i | \text{either } a'_i = a_i \text{ or } \pi_i(t_i, a'_i, a_{-i}) > \pi_i(t_i, a_i, a_{-i})\}.
\]

Let \(R_{G(t_1, \ldots, t_N)} : A \rightarrow A\) be the same-or-better-reply correspondence of the game \(G(t_1, \ldots, t_N)\), i.e.,

\[
R_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) = \prod_{i \in N} R_{t_i}(a_i, a_{-i}).
\]

**Definition 3.1.** A set of action profiles \(\hat{A} \subseteq A\) in \(G(t_1, \ldots, t_N)\) is closed under same-or-better replies (henceforth, CUSOBR set) if for all \((a_1, \ldots, a_N) \in \hat{A}\), we have \(R_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) \subseteq \hat{A}\). A minimal CUSOBR set is a CUSOBR set that does not contain a proper subset that is a CUSOBR set.

For any nonempty \(\hat{A} \subseteq A\), define

\[
\tilde{R}_{G(t_1, \ldots, t_N)}(\hat{A}) = \bigcup_{(a_1, \ldots, a_N) \in \hat{A}} \left( \prod_{i \in N} R_{t_i}(a_i, a_{-i}) \right).
\]

Equivalently, \(\hat{A}\) is a CUSOBR set if and only if \(\hat{A}\) is a fixed point of \(\tilde{R}_{G(t_1, \ldots, t_N)}\), i.e.,

\[
\tilde{R}_{G(t_1, \ldots, t_N)}(\hat{A}) = \hat{A}.
\]

It is easy to see that \((a_1, \ldots, a_N)\) is a pure Nash equilibrium of \(G(t_1, \ldots, t_N)\) if and only if \(\{(a_1, \ldots, a_N)\}\) is a singleton minimal CUSOBR set. Furthermore, since the game \(G(t_1, \ldots, t_N)\) has a finite number of action profiles, there exists a minimal CUSOBR set.

CUSOBR sets are related to product sets that are closed under weakly better replies (Ritzberger and Weibull, 1995); we discuss the differences at present. For any \((a_i, a_{-i}) \in A\), the set of weakly-better replies for the player of type \(t_i\) is

\[
WBR_{t_i}(a_i, a_{-i}) = \{a'_i \in A_i | \pi_i(t_i, a'_i, a_{-i}) \geq \pi_i(t_i, a_i, a_{-i})\}.
\]

Let \(WBR_{G(t_1, \ldots, t_N)} : A \rightarrow A\) be the weakly-better-reply correspondence of the game
\[
\begin{array}{ccc}
U & L & R \\
D & 2.0 & 3.2 \\
\end{array}
\quad
\begin{array}{ccc}
U & L & R \\
D & 2.0 & 3.2 \\
\end{array}
\]

(a) (b)

\textbf{Figure 1:} Relation between minimal CUSOBR and minimal PCUWBR sets

For any \( G(t_1, \ldots, t_N) \), i.e.,

\[ WBR_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) = \prod_{i \in N} WBR_{G}(a_i, a_{-i}). \]

A \textit{product set} of action profiles \( \hat{A} \subseteq A \) is such that \( \hat{A} = \prod_{i \in N} \hat{A}_i \), where \( \emptyset \neq \hat{A}_i \subseteq A_i, \forall i \in N \). Then, \( \hat{A} \) is a \textit{product set of action profiles that is closed under weakly better replies} (or PCUWBR set) if \( \hat{A} \) is a product set and \( WBR_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) \subseteq \hat{A} \) for all \((a_1, \ldots, a_N) \in \hat{A}\).\(^{15}\) A \textit{minimal PCUWBR set} is a PCUWBR set that does not contain a proper subset that is a PCUWBR set.

Since \( R_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) \subseteq WBR_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) \), every PCUWBR set is a CUSOBR set. Hence, every minimal PCUWBR set contains a minimal CUSOBR set. Therefore, in some games, the set of minimal CUSOBR sets is a refinement of the set of minimal PCUWBR sets. Game (a) in Figure 1 is an example; its unique minimal CUSOBR set \{\((D, R)\)\} is a refinement of its unique minimal PCUWBR set \{\((U, L), (U, R), (D, L), (D, R)\)\}. However, it is not necessary that every minimal CUSOBR set of a game is a subset of some minimal PCUWBR set. Game (b) in Figure 1 is an example; \{\((D, L)\)\} is its unique minimal PCUWBR set whereas it has two minimal CUSOBR sets, \{\((D, L)\)\} and \{\((U, R)\)\}. Finally, unlike minimal PCUWBR sets, minimal CUSOBR sets are not necessarily product sets (for example, see the case when \( t_i = t_{-i} = II \) in Figure 2).

Consider any fixed matching in which \( \{t_1, \ldots, t_N\} \) are the types of the selected players. Define a \textit{state} of the matched players in a period as the current action profile \((a_1, \ldots, a_N)\) of the players. Hence, \( A \) is the set of states of the Markov process.

\(^{15}\)This definition is equivalent to the original definition by Ritzberger and Weibull (1995) where the weakly-better-reply correspondence is defined over the domain of mixed strategies. For any \( \hat{A}_i \subseteq A_i \), let \( S(\hat{A}_i) \) be the set of mixed strategies with support in \( \hat{A}_i \). The weakly-better-reply correspondence over the domain of mixed strategies, \( WBR_{G(t_1, \ldots, t_N)} : \prod_{i \in N} S(\hat{A}_i) \rightarrow \hat{A} \) is such that \( WBR_{G(t_1, \ldots, t_N)}(s_1, \ldots, s_N) = \prod_{i \in N} WBR_{G}(s_i, a_{-i}) \), where \( WBR_{G}(s_i, a_{-i}) = \{a'_i \in A_i | E_{s_i, a_{-i}}(\pi_{i}(t_i, a'_i, a_{-i})) \geq E_{s_i, a_{-i}}(\pi_{i}(t_i, a_i, a_{-i}))\}, \forall i \in N \). The equivalence follows since for any product set \( \hat{A} = \prod_{i \in N} \hat{A}_i \), \( WBR_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) \subseteq \hat{A}, \forall(a_1, \ldots, a_N) \in \hat{A} \iff WBR_{G(t_1, \ldots, t_N)}(s_1, \ldots, s_N) \subseteq \hat{A}, \forall(s_1, \ldots, s_N) \in \prod_{i \in N} S(\hat{A}_i) \).
$\mathcal{M}_{FM/G}(q)$. Our interest in minimal CUSOBR sets stems from the following result:

**Proposition 3.2.** Let $\{t_1, \ldots, t_N\}$ be the types of the matched players under fixed matching. Then, $\hat{A} \subseteq A$ is a recurrent class of $\mathcal{M}_{FM/G}(q)$ if and only if $\hat{A}$ is a minimal CUSOBR set of $G(t_1, \ldots, t_N)$.\(^{16}\)

Thus, under fixed matching, when players adapt their actions using the ex-post regret heuristic rule, per-period play will almost surely enter a minimal CUSOBR set – the specific set depends on the initial condition – and then stay inside this set forever, with every action profile in this set being played infinitely often.

### 3.1 Weakly Acyclic Games

We can establish a stronger result if $G(t_1, \ldots, t_N)$ is weakly acyclic under better replies. A **better-reply graph** is defined as follows: each action profile of $G(t_1, \ldots, t_N)$ is a vertex of the graph and there exists a directed edge from vertex $(a_1, \ldots, a_N)$ to vertex $(a'_1, \ldots, a'_N)$ if and only if $(a_1, \ldots, a_N) \neq (a'_1, \ldots, a'_N)$ and $(a'_1, \ldots, a'_N) \in R_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N)$. A **sink** is a vertex with no outgoing edges. A **better-reply path** is a sequence of vertices $(a_1^1, \ldots, a_N^L)$ such that there exists a directed edge from each $(a_1^l, \ldots, a_N^l)$ to $(a_1^{l+1}, \ldots, a_N^{l+1})$. The game $G(t_1, \ldots, t_N)$ is **weakly acyclic under better replies** if from any action profile, there exists at least one better-reply path to a sink.\(^{17}\) Clearly, an action profile is a sink if and only if it is a pure Nash equilibrium of $G(t_1, \ldots, t_N)$. Thus, the game $G(t_1, \ldots, t_N)$ is weakly acyclic under better replies if from any action profile there exists at least one better-reply path to a pure Nash equilibrium.

It is straightforward to show that if $G(t_1, \ldots, t_N)$ is weakly acyclic under better replies, then every CUSOBR set of $G(t_1, \ldots, t_N)$ must contain a pure Nash equilibrium of $G(t_1, \ldots, t_N)$. Since a pure Nash equilibrium is a minimal CUSOBR set, it follows that if $G(t_1, \ldots, t_N)$ is weakly acyclic under better replies, then the set of its minimal CUSOBR sets coincides with the set of its pure Nash equilibria. Therefore, we easily obtain the following corollary from Proposition 3.2:

\(^{16}\)This result, however, does not hold when players have longer than one-period memory. In that case, any recurrent class is an $\omega$-set, i.e., a CUSOBR set that does not contain any product set which is also a CUSOBR set. See Saran and Serrano (2012) for details.

\(^{17}\)Young (2004) defines a class of weakly acyclic complete information games that is a subset of the class of games that are weakly acyclic under better replies. Ordinal potential games and supermodular games are also weakly acyclic under better replies. See Saran and Serrano (2012) for details.
Corollary 3.3. Let \( \{t_1, \ldots, t_N\} \) be the types of the matched players under fixed matching and suppose \( G(t_1, \ldots, t_N) \) is weakly acyclic under better replies. Then, \( \hat{A} \) is a recurrent class of \( \mathcal{M}_{FM/G}(q) \) if and only if \( \hat{A} = \{(a_1, \ldots, a_N)\} \) such that \( (a_1, \ldots, a_N) \) is a pure Nash equilibrium of \( G(t_1, \ldots, t_N) \).

Thus, in spite of the initial incomplete information, the ex-post regret heuristic rule under fixed matching gives players enough experience so that, when switches happen in the direction of ex-post regrets and the game is weakly acyclic under better replies, per-period play in the game almost surely converges to a pure-strategy Nash equilibrium of the complete information game corresponding to the types of the players in the match.

3.2 Applications

We now illustrate our results by studying specific games under fixed matching.

**Fashion Game:** Consider the following incomplete information version of the Fashion Game (Young, 1998). There are two positions and for each position there are two types of players. Type \( I \) is a conformist who prefers to wear the same color as the other player. Type \( II \), on the other hand, is a nonconformist who likes to wear a color that contrasts best with that worn by the other player. If the other player wears blue, type \( II \) prefers to wear yellow; if the other player wears yellow, type \( II \) prefers to wear red; and finally, if the other player wears red, type \( II \) prefers to wear blue. Let \( b, y, \) and \( r \) stand for blue, yellow and red, respectively. Figure 2 shows the different complete information games \( G(t_i, t_{-i}) \) corresponding to various realizations of types of players \( i \) and \(-i\).

First, consider the case when \( t_i = t_{-i} = I \), i.e., both players are conformist. In this case, \( G(t_i, t_{-i}) \) is weakly acyclic under better replies. The game has three pure Nash equilibria \( (b, b) \), \( (y, y) \), and \( (r, r) \). Thus, in the fixed matching model, if the realized types of the two players are such that both players are conformist, then it follows from Corollary 3.3 that per-period play will converge almost surely in finite time to one of these pure Nash equilibria – the specific equilibrium depends on the initial condition.

---

18Our dynamics is uncoupled in the sense of Hart and Mas-Colell (2003) and convergence to Nash equilibrium is obtained only in a specific class of games, while those authors seek a convergence result in all games. Our convergence to Nash equilibrium is a consequence of the fixed matching and the weakly acyclic property of the game. Saran and Serrano (2012) show that this result also holds when players have longer than one-period memory.
Second, consider the case when \( t_i = t_{-i} = I \), i.e., both players are nonconformist. In this case, \( G(t_i, t_{-i}) \) is has a unique minimal CUSOBR set equal to \( \{(y, b), (y, r), (b, r), (b, y), (r, y), (r, b)\} \). Thus, in the fixed matching model, if the realized types of the two players are such that both players are nonconformist, then it follows from Proposition 3.2 that per-period play will almost surely enter this unique minimal CUSOBR set in finite time. After that point, per-period play will cycle in a fixed order: each spell of \((y, b)\) will be followed by spells of \((y, r), (b, r), (b, y), (r, y), (r, b)\), in exactly that order, before play cycles back to \((y, b)\) – the length of each spell is stochastic because of positive probability of inertia.

Last, consider the case when \( t_i = II \) and \( t_{-i} = I \), i.e., player \( i \) is nonconformist whereas player \(-i\) is conformist. In this case, \( G(t_i, t_{-i}) \) is has a unique minimal CUSOBR set equal to \( \{(b, b), (y, b), (y, y), (r, y), (r, r), (b, r)\} \). Thus, in the fixed matching model, if the realized types are such that player \( i \) is nonconformist whereas player \(-i\) is conformist, then it follows from Proposition 3.2 that per-period play will almost surely enter this unique minimal CUSOBR set in finite time. After that point, per-period play will cycle in a fixed order: each spell of \((b, b)\) will be followed by spells of \((y, b), (y, y), (r, y), (r, r), (b, r)\), in exactly that order, before play cycles back to \((b, b)\) – again, the length of each spell is stochastic. \(\diamond\)

**Second-Price Auction:** Consider the second-price auction with \( N \geq 2 \) positions. The type of any player in position \( i \) is equal to her valuation \( v_i \) for an object, where \( v_i \in Z \equiv \{0, \delta, 2\delta, \ldots, 1\} \). In this game, players simultaneously announce their bids for the object, which can be any number in \( Z \). The player with the highest bid wins
the object and pays a price equal to the second-highest bid.\footnote{If more than one player bid the highest amount, then the object is allocated at random among the highest bidders and the price is equal to their bid. We make the same tie-breaking assumption in the first-price auction.}

If the object is traded at price \( p \), the winning player \( i \)'s payoff is \( v_i - p \) and all the other players get a payoff of 0.

Pick any profile of valuations \( \{v_1, \ldots, v_N\} \) and consider the second-price auction when these valuations are common knowledge. We have the following result:

**Lemma 3.4.** The one-shot second-price auction in which the valuations \( \{v_1, \ldots, v_N\} \) of the players are common knowledge is weakly acyclic under better replies.

In the fixed matching model, the selected players repeatedly play the second-price auction. Let \( \mathcal{M}_{FM/SPA}(q) \) denote the Markov process when the players use the ex-post regret heuristic rule to adapt their bids in the second-price auction under fixed matching. The set of states of this Markov process equals \( Z^N \). The following corollary follows from the above lemma and Corollary 3.3:

**Corollary 3.5.** Let \( \{v_1, \ldots, v_N\} \) be the valuations of the matched players under fixed matching when the game is the second-price auction. Then, \( \hat{A} \subseteq Z^N \) is a recurrent class of \( \mathcal{M}_{FM/SPA}(q) \) if and only if \( \hat{A} = \{(a_1, \ldots, a_N)\} \) such that \( (a_1, \ldots, a_N) \) is a pure Nash equilibrium of the one-shot second-price auction in which the valuations \( \{v_1, \ldots, v_N\} \) of the players are common knowledge.\footnote{Strictly speaking, we cannot immediately apply Corollary 3.3 here because of the difference between the definitions of payoffs in the game and payoffs in the ex-post regret heuristic rule (see Footnote 12). Nevertheless, in a second-price auction, the expected payoff of valuation \( v_i \) from bidding \( a'_i \) against \( a_{-i} \) is greater than her expected payoff from bidding \( a_i \) against \( a_{-i} \) if and only if the expected payoff of valuation \( v_i \) from bidding \( a'_i \) against \( a_{-i} \) is greater than at least one of her realized payoffs from bidding \( a_i \) against \( a_{-i} \). This fact and Lemma 3.4 are sufficient to establish the corollary. Similar comments are applicable to Corollaries 3.7, 3.9, 4.6, and when we apply Lemma 4.7 to \( \mathcal{M}_{RM/FPA}(q) \) and Lemma 4.8 to \( \mathcal{M}_{RM/BD}(q) \).} \( \diamond \)

**First-Price Auction:** Now suppose that the game is the first-price auction with \( N \geq 2 \) positions. The only change in the rules is that the player who wins the object pays a price equal to her bid.

**Lemma 3.6.** The one-shot first-price auction in which the valuations \( \{v_1, \ldots, v_N\} \) of the players are common knowledge is weakly acyclic under better replies.

Letting \( \mathcal{M}_{FM/FPA}(q) \) denote the Markov process under fixed matching when the game is the first-price auction, we obtain the following corollary using the above lemma and Corollary 3.3:
Corollary 3.7. Let \(\{v_1, \ldots, v_N\}\) be the valuations of the matched players under fixed matching when the game is the first-price auction. Then, \(\hat{A} \subseteq \mathbb{Z}^N\) is a recurrent class of \(\mathcal{M}_{FM/FPA}(q)\) if and only if \(\hat{A} = \{(a_1, \ldots, a_N)\}\) such that \((a_1, \ldots, a_N)\) is a pure Nash equilibrium of the one-shot first-price auction in which the valuations \(\{v_1, \ldots, v_N\}\) of the players are common knowledge.

\[\diamond\]

**Bertrand Duopoly:** Consider next the game of Bertrand duopoly. This game has \(N = 2\) positions. The type of a player in position \(i\) is equal to her constant marginal cost \(c_i\) of producing a good, where \(c_i \in C \equiv \{0, \delta, 2\delta, \ldots, x\}\) (note that \(x\) is some multiple of \(\delta\)). The market demand for the good at price \(p\) is \(Q = \frac{x - p}{y}\), where \(x, y > 0\). In this game, the two players simultaneously post their prices \(p_1\) and \(p_2\) for the object, where each \(p_i \in P = \{0, \frac{\delta}{2}, \delta, \ldots, x - \frac{\delta}{2}, x\}\) (this ensures that each marginal cost type \(c_i\) can post her monopoly profit maximizing price \(p^*(c_i) = \frac{x + c_i}{2}\)).

The player who posts a lower price captures the market, i.e., supplies the entire market demand. If both players post the same price, then each player captures the market with equal probability. Hence, for any \((p_1, p_2)\), player \(i\)’s realized payoff is 0 if either \(p_i > p_j\) or \(p_i = p_j\) and she loses the tie-breaking lottery, and \(\frac{(p_i - c_i)(x - p_i)}{y}\) if either \(p_i < p_j\) or \(p_i = p_j\) and she wins the tie-breaking lottery.\(^{21}\)

Pick any profile of marginal costs \(\{c_1, c_2\}\) and consider the Bertrand duopoly game in which these marginal costs are common knowledge. We have the following result:

**Lemma 3.8.** The one-shot Bertrand duopoly game in which the marginal costs \(\{c_1, c_2\}\) of the players are common knowledge is weakly acyclic under better replies.\(^{22}\)

In the fixed matching model, the selected players repeatedly play the Bertrand duopoly game. Let \(\mathcal{M}_{FM/BD}\) be the Markov process when the players use the ex-post regret heuristic rule to adapt their posted prices in the Bertrand duopoly game under fixed matching. The set of states of this Markov process equals \(P^2\). The following corollary follows from the above lemma and Corollary 3.3:

**Corollary 3.9.** Let \(\{c_1, c_2\}\) be the marginal costs of the matched players under fixed matching when the game is Bertrand duopoly. Then, \(\hat{A} \subseteq P^2\) is a recurrent class of \(\mathcal{M}_{FM/BD}(q)\) if and only if \(\hat{A} = \{(p_1, p_2)\}\) such that \((p_1, p_2)\) is a pure Nash equilibrium of the one-shot Bertrand duopoly game in which the marginal costs \(\{c_1, c_2\}\) of the players are common knowledge.

\[^{21}\text{See Spulber (1995) for the equilibrium analysis of the winner-takes-all Bertrand competition with unknown but continuously distributed marginal costs.}\]

\[^{22}\text{The proof of this and subsequent results for Bertrand duopoly are quite close to those for the first-price auction, and hence we omit them (available upon request).}\]
Thus in the second-price and first-price auctions and Bertrand duopoly game, the dynamics generated by our ex-post regret heuristic rule converges to one of the pure Nash equilibria of the game; which equilibrium will be the limit of the dynamics is a function of the initial condition. The general point, nonetheless, is how Nash play emerges through using the heuristic rule on the fixed set of players, in spite of the restrictive informational assumptions.

4 Long-Run Behavior under Random Matching

Reconsider the incomplete information game $G$. We let the strategy of type $t_i$, denoted by $\sigma_i(t_i)$, be a distribution of actions $A_i$ in population $P_i$.\(^ {23}\) Given $\sigma_i(t_i)$, let $\sigma_i(t_i, a_i)$ be the relative frequency of action $a_i$, and $A_i(\sigma_i(t_i))$ denote the support of $\sigma_i(t_i)$. The strategy of position $i$, denoted by $\sigma_i$, is a collection of strategies of all types $t_i \in T_i$. Let $\Sigma_i$ be the set of strategies of position $i$. Let $\Sigma = \prod_{i \in N} \Sigma_i$ be the set of strategy profiles. Some of the following definitions are also found in Saran and Serrano (2014).

We say that strategy $\sigma_i'$ is adjacent to strategy $\sigma_i$ if there exists exactly one type $t_i^* \in T_i$ such that $\sigma_i'(t_i^*) \neq \sigma_i(t_i^*)$ and we can obtain $\sigma_i'(t_i^*)$ from $\sigma_i(t_i^*)$ by letting exactly one player in $P_i$ change her action from some $a_i^{1*} \in A_i(\sigma_i(t_i^*))$ to some $a_i^{2*} \in A_i(\sigma_i'(t_i^*))$. For any $\sigma_i \in \Sigma_i$, let

$$\Sigma_i(\sigma_i) = \{\sigma_i' \in \Sigma_i | \text{either } \sigma_i' = \sigma_i \text{ or } \sigma_i' \text{ is adjacent to } \sigma_i\}.$$

The same or adjacent ex-post better reply correspondence $R_G : \Sigma \rightarrow \Sigma$ be such that $(\sigma_1', \ldots, \sigma_N') \in R_G(\sigma_1, \ldots, \sigma_N)$ if and only if

1. $\sigma_i' \in \Sigma_i(\sigma_i), \forall i \in N$, and
2. if $(\sigma_1', \ldots, \sigma_N') \neq (\sigma_1, \ldots, \sigma_N)$, then letting $I = \{i | \sigma_i \neq \sigma_i'\}$, there must exist $(t_1, \ldots, t_N)$ and $(a_1, \ldots, a_N) \in \prod_{i \in N} A_i(\sigma_i(t_i))$ such that for all $i \in I$, we have $t_i = t_i^*$, $a_i = a_i^{1*}$ and $$\pi_i(t_i^*, a_i^{2*}, a_{-i}) > \pi_i(t_i^*, a_i^{1*}, a_{-i}).$$

\(^ {23}\)We are abusing the standard definition since a strategy of type $t_i$ in game $G$ should include any probability distribution over the set of actions $A_i$. We are considering only that subset of these probability distributions in which the probability of any action is equal to $\frac{1}{|P_i|}$ for some nonnegative integer $x \leq |P_i|$. 

17
According to the first condition, \((\sigma'_1, \ldots, \sigma'_N)\) must be such that \(\sigma'_i\) is either the same or adjacent to \(\sigma_i\) for all \(i \in N\). If \((\sigma'_1, \ldots, \sigma'_N)\) differs from \((\sigma_1, \ldots, \sigma_N)\), then it follows from the definition of \(I\) that \(\sigma'_i\) is adjacent to \(\sigma_i\) for all \(i \in I\). The second condition says that for all \(i \in I\), this change in strategy should be justified as an ex-post better reply, i.e., there must exist an ex-post event in which for all \(i \in I\), the realized type \(t_i = t^*_i\), the realized action \(a_i = a^*_i\), and the action \(a^*_{i+1}\) is a better reply to \(a_{i+1}\) than the action \(a^*_{i-1}\).

**Definition 4.1.** A set of strategy profiles \(\hat{\Sigma} \subseteq \Sigma\) in \(G\) is **closed under same-or-adjacent-ex-post-better replies** (henceforth, CUSOAEBR set) if for all \((\sigma_1, \ldots, \sigma_N) \in \hat{\Sigma}\), we have \(R_G(\sigma_1, \ldots, \sigma_N) \subseteq \hat{\Sigma}\). A **minimal CUSOAEBR set** is a CUSOAEBR set that does not contain a proper subset that is a CUSOAEBR set.

For any nonempty \(\hat{\Sigma} \subseteq \Sigma\), define

\[
\hat{R}_G(\hat{\Sigma}) = \bigcup_{(\sigma_1, \ldots, \sigma_N) \in \hat{\Sigma}} R_G(\sigma_1, \ldots, \sigma_N).
\]

Equivalently, \(\hat{\Sigma}\) is a CUSOAEBR set if and only if \(\hat{\Sigma}\) is a fixed point of \(\hat{R}_G\), i.e., \(\hat{R}_G(\hat{\Sigma}) = \hat{\Sigma}\). Since \(G\) has a finite number of strategy profiles, there exists a minimal CUSOAEBR set.

**Definition 4.2.** A strategy profile \((\sigma_1, \ldots, \sigma_N) \in \Sigma\) is a **pure ex-post equilibrium** of \(G\) if for all \(i \in N\), \((t_i, t_{-i}), (a_i, a_{-i}) \in \prod_{j \in N} A_j(\sigma_j(t_j))\),

\[
\pi_i(t_i, a_i, a_{-i}) \geq \pi_i(t_i, a'_i, a_{-i}), \forall a'_i \in A_i. \tag{24}
\]

The following lemma shows that every pure ex-post equilibrium is equivalent a singleton CUSOAEBR set.

**Lemma 4.3.** A strategy profile \((\sigma_1, \ldots, \sigma_N) \in \Sigma\) is a pure ex-post equilibrium of \(G\) if and only if \(\{(\sigma_1, \ldots, \sigma_N)\}\) is a singleton CUSOAEBR set.

Consider the Markov process \(M_{RM/G}(q)\) generated by random matching of the players and our ex-post regret heuristic rule. A state is a list specifying for all \(i \in N\)...

---

24There is a slight abuse of language in our use of the term “pure” here. We are in fact allowing for mixed strategies under the condition that in every ex-post event, every action profile that is in the support of the mixed-strategy profile corresponding to that ex-post event is a pure Nash equilibrium. But there is a purification argument for any such mixed-strategy profile, by having the appropriate proportions in the population of each type play the pure actions in the support of the mixed strategy.
and \( t_i \in T_i \), the distribution of actions \( A_i \) in population \( \mathcal{P}_i \). Hence, a state is a strategy profile of \( G \) and the set of states is \( \Sigma \). We characterize the recurrent classes of the Markov process \( \mathcal{M}_{RM/G}(q) \) in the next result.

**Proposition 4.4.** \( \hat{\Sigma} \subseteq \Sigma \) is a recurrent class of \( \mathcal{M}_{RM/G}(q) \) if and only if \( \hat{\Sigma} \) is a minimal CUSOAEBR set of \( G \).

### 4.1 Applications

We now illustrate our results in specific games.

**Second-Price Auction:** Let \( G \) be the second-price auction with incomplete information in which there are \( N \geq 2 \) positions, and the set of valuations and the set of bids for each position are equal to \( Z \). The next lemma shows that every minimal CUSOAEBR set of the second-price auction is singleton.

**Lemma 4.5.** If \( \hat{\Sigma} \subseteq \Sigma \) is a minimal CUSOAEBR set of the second-price auction, then \( \hat{\Sigma} \) is singleton.\(^{25}\)

Let \( \mathcal{M}_{RM/SPA}(q) \) be the Markov process under random matching when the game is the second-price auction. We have the following result using Lemmas 4.3 and 4.5, and Proposition 4.4:

**Corollary 4.6.** In the random-matching model, when the game is second-price auction, \( \hat{\Sigma} \subseteq \Sigma \) is a recurrent classes of \( \mathcal{M}_{RM/SPA}(q) \) if and only if \( \hat{\Sigma} = \{(\sigma_1, \ldots, \sigma_N)\} \) such that \( (\sigma_1, \ldots, \sigma_N) \) is a pure ex-post equilibrium of the second-price auction.

**Remark:** One pure ex-post equilibrium of the second-price auction is truth telling. However, there also exist pure ex-post equilibria that are ex-post inefficient (this is also true for the standard second-price auction model with continuous set of valuations). An example is the following strategy profile: let \( n \in (0, 1/\delta) \) and

\[
\sigma_1(v_1, n\delta) = 1, \forall v_1 \leq n\delta, \quad \text{and} \quad \sigma_1(v_1, v_1) = 1, \forall v_1 > n\delta,
\]

and for all \( i \neq 1 \)

\[
\sigma_i(v_i, 0) = 1, \forall v_i \leq n\delta, \quad \text{and} \quad \sigma_i(v_i, v_i) = 1, \forall v_i > n\delta.
\]

\(^{25}\)Incomplete information games like the second-price auction in which all minimal CUSOAEBR sets are singleton are weakly acyclic under the adjacent ex-post better reply operation in the sense that for any strategy profile \( \sigma \), there exists a path \( (\sigma^1, \ldots, \sigma^L) \) such that \( \sigma^1 = \sigma, \sigma^{l+1} \in R_G(\sigma^l), \forall l \), and \( \sigma^L \) is a pure ex-post equilibrium of the game.
Hence, even such inefficient pure ex-post equilibria are absorbing states of the Markov process under random matching.\textsuperscript{26}

Ex-post equilibria rarely exist in games of incomplete information – the second-price auction is special in that regard. As a result, minimal CUSOAEBR sets are usually not singleton. In fact, minimal CUSOAEBR sets are typically very large, offering limited prediction. This is illustrated below for the first-price auction and Bertrand duopoly.

**First-Price Auction:** Let $G$ be the first-price auction with incomplete information in which there are $N \geq 2$ positions, and the set of valuations and the set of bids for each position are equal to $Z$.

**Lemma 4.7.** Suppose $2\delta < 1$. The first-price auction has a unique minimal CUSOAEBR set $\hat{\Sigma}$. Furthermore, $\forall i \in N, v_i > 2\delta$, and $a_i \in [\delta, v_i - 2\delta] \cap Z$, there exists $(\sigma_1, \ldots, \sigma_N) \in \hat{\Sigma}$ such that $\sigma_i(v_i, a_i) = 1$.

Let $\mathcal{M}_{RM/FPA}(q)$ be the Markov process under random matching when the game is first-price auction. As a corollary of the above lemma and Proposition 4.4, it follows that when $\delta$ is small enough, the process $\mathcal{M}_{RM/FPA}(q)$ has a unique large recurrent class. For each $v_i > 2\delta$ and each bid $a_i \in [\delta, v_i - 2\delta]$, we can find a strategy profile in this recurrent class such that all players of valuation $v_i$ bid $a_i$. Hence, looking at this recurrent class is not informative.

**Bertrand Duopoly:** Let $G$ now be the Bertrand duopoly with incomplete information in which there are 2 positions, the set of constant marginal costs equal to $C$, and the set of prices equal to $P$ for each position.

**Lemma 4.8.** The Bertrand duopoly game has a unique minimal CUSOAEBR set $\hat{\Sigma}$. Furthermore, $\forall i \in \{1, 2\}, c_i \in C \setminus \{0, x\}$, and $p_i \in [c_i + \frac{\delta}{2}, p^*(c_i)] \cap P$, there exists $(\sigma_1, \sigma_2) \in \hat{\Sigma}$ such that $\sigma_i(c_i, p_i) = 1$.

Let $\mathcal{M}_{RM/BD}(q)$ be the Markov process under random matching when the game is Bertrand duopoly. As a corollary of the above lemma and Proposition 4.4, it follows that the process $\mathcal{M}_{RM/BD}(q)$ has a unique large recurrent class. Again, looking at this recurrent class is not informative.

\textsuperscript{26}Note how if we had a single population model – which is a reasonable alternative model since the game is symmetric –, these inefficient equilibria would be eliminated. Consider the above strategy profile. With a single population, the players in position 1 will also meet each other, and hence, those who bid above their valuations will experience positive regret. The assumption of single versus multiple populations however does not make any qualitative difference in first-price auction and Bertrand duopoly.
5 Long-Run Behavior under Random Matching when Players make Mistakes

As demonstrated above for first-price auction and Bertrand duopoly games, the recurrent classes of the Markov process generated by our ex-post heuristic rule under random matching are typically very large. In an attempt to provide a sharper prediction in these cases, we shall now specify a perturbed version of the ex-post heuristic rule. Specifically, we make two changes:\(^\text{27}\)

- We fix \(q(\Delta_i(t_i, a_i, a'_i, a_{-i})) = \frac{1}{a_i(t_i, a_i, a'_i, a_{-i})},\) i.e., \(q\) is now a positive parameter that is sufficiently small to ensure positive probability of inertia. Thus switches to other actions are an increasing function of the corresponding ex-post regret, and the way in which ex-post regret affects switching probabilities is exponential.

- Additionally, we assume that if \(\Delta_i(t_i, a_i, a'_i, a_{-i}) \leq 0,\) then player \(i\) changes her action to \(a'_i\) in period \(l+1\) with probability \(q^{\frac{1}{\gamma}},\) where \(\gamma > 0\) is smaller than any positive ex-post regret.\(^\text{28}\) Allowing for a positive \(\gamma\) leads to ergodic dynamics.

Fix \(\alpha \equiv \frac{1}{q^{\frac{1}{\gamma}}}\) to be the probability of an individual’s “mistake” – a switch under no regret. Let \(\mathcal{M}_{RM/G}(q, \alpha)\) be the stochastic process generated by the above perturbed ex-post heuristic rule under random matching. In Saran and Serrano (2014), we study this process when \(G\) is a \(2 \times 2\) game of incomplete information. As explained there, the standard procedure of taking the mistake probability \(\alpha\) to zero (which is the same as taking \(\gamma\) to zero when \(q\) is fixed at some value less than 1) is not useful. This exercise can only select amongst the recurrent classes of \(\mathcal{M}_{RM/G}(q, 0)\) – the Markov process generated by the ex-post heuristic rule of Section 2 under random matching –, which proves to be fruitless especially in cases like first-price auction and Bertrand duopoly where \(\mathcal{M}_{RM/G}(q, 0)\) has a unique large recurrent class.

As in Saran and Serrano (2014), we instead fix \(\gamma > 0\) and take the limit with respect to \(q \rightarrow 0.\) By doing so, the process \(\mathcal{M}_{RM/G}(q, \alpha)\) becomes a regular perturbation of \(\mathcal{M}_{RM/G}(0, 0).\)\(^\text{29}\) Hence, as \(q \rightarrow 0,\) the support of the unique invariant distribution of \(\mathcal{M}_{RM/G}(q, \alpha)\) is contained within the recurrent classes of \(\mathcal{M}_{RM/G}(0, 0),\)

\(^{27}\)The results of this section are valid for other specifications of switching probabilities. See Section 2.2 in Saran and Serrano (2014).

\(^{28}\)This minimum positive ex-post regret is well-defined since the sets of players, types and actions are finite.

\(^{29}\)See the Appendix for the definition of regular perturbation and other concepts related to finite Markov processes and stochastic stability.
which are all singletons – in fact, since the process $M_{RM/G}(0,0)$ is static, the set of absorbing states of $M_{RM/G}(0,0)$ coincides with its set of states. Following Kandori, Mailath and Rob (1993) and Young (1993), we refer to the states that are in the support of this limiting distribution as stochastically stable. The limiting distribution approximates both the frequency with which a state is visited over a long horizon and the probability of being in a particular state at a point in time $l$ when $l$ is sufficiently large. Hence, the stochastically stable states are the only states on which the system will spend a positive proportion of time in the very long run when the switches of actions are possible but very unlikely events.

For any game $G$, define the weighted graph $W_G$ with set of vertices equal to $\Sigma$ as follows: there exists a directed edge from strategy profile $\sigma$ to $\sigma'$ if $\sigma \neq \sigma'$ and $\sigma' \in R_G(\sigma)$, and the weight of this directed edge is equal to the resistance $r(\sigma, \sigma')$.\(^{30}\) A weighted cycle is a directed path on this graph that begins and ends at the same vertex with no other repetition of vertices. Saran and Serrano (2014) prove that if a state passes the weighted cycle test given below, then it is not stochastically stable.

**Proposition 5.1.** (Weighted cycle test) $\sigma$ is not stochastically stable if in the weighted graph $W_G$, there exists an outgoing edge from $\sigma$ with weight less than $\frac{1}{\gamma}$ and either of the following holds:

1. There does not exist any weighted cycle containing this outgoing edge.
2. In all weighted cycles containing this outgoing edge, the weight of this outgoing edge is less than the weight of some other edge on the weighted cycle.\(^{31}\)

### 5.1 Applications

We now apply the weighted cycle test to specific games. As we will see, the analysis using stochastic stability points in the direction of minimax regret equilibrium (see Saran and Serrano (2014) for more details on minimax regret equilibria). For any $t_i$,

\(^{30}\)See the Appendix for the definition of resistance. To calculate $r(\sigma, \sigma')$, consider any match that can occur in state $\sigma$, which includes all the players who switch their actions between $\sigma$ and $\sigma'$. For each such match, sum the reciprocals of the ex-post regrets of the players who switch their actions between $\sigma$ and $\sigma'$ (if the ex-post regret of any player is nonpositive, then add $\frac{1}{\gamma}$ instead). The lowest such sum is $r(\sigma, \sigma')$.

\(^{31}\)We cannot use the weighted cycle test to claim that if $\sigma$ is stochastically stable, then $\sigma$ belongs to a minimal CUSOAEBR set. To see this, let $\sigma^1$ be stochastically stable and not belong to any minimal CUSOAEBR set. We know that there exists a directed path $\sigma^1, \ldots, \sigma^L$ such that $\sigma^L$ is in some minimal CUSOAEBR set. We also know that there does not exist any weighted cycle containing this path. However, there could exist a weighted cycle containing the first edge $\sigma^1, \sigma^2$.\(^{55}\)
\(a_i, a_i' \neq a_i,\) and \(a_{-i},\) define
\[
\tilde{\Delta}_i(t_i, a_i, a_i', a_{-i}) = \max\{\Delta_i(t_i, a_i, a_i', a_{-i}), 0\}.
\]
Action \(a_i\) is a minimax regret best response to \(\sigma_{-i}\) for type \(t_i\) if
\[
a_i \in \arg\min_{\hat{a}_i \in A_i} \left( \max_{a_{-i} \in \Pi_{j \neq i} (\bigcup_{t_j \in T_j} A_j(t_j))} \left( \max_{a_i' \neq \hat{a}_i} \tilde{\Delta}_i(t_i, \hat{a}_i, a_i', a_{-i}) \right) \right).
\]

**Definition 5.2.** \(\sigma\) is a minimax regret equilibrium if for all \(i \in N\) and \(t_i \in T_i\), every action \(a_i \in A_i(\sigma_i(t_i))\) is a minimax regret best response to \(\sigma_{-i}\) for type \(t_i\).

**Second-Price Auction:** Let \(M_{RM/SPA}(q, \alpha)\) be the Markov process with mistakes under random matching when the game is second-price auction. We have the following result:

**Proposition 5.3.** In the random-matching model, when the game is second-price auction, if \((\sigma_1, \ldots, \sigma_N) \in \Sigma\) is stochastically stable in \(M_{RM/SPA}(q, \alpha)\) as \(q \to 0\), then \((\sigma_1, \ldots, \sigma_N)\) is a pure ex-post equilibrium of the second-price auction.

Intuitively, if \(\sigma\) is not a pure ex-post equilibrium of the second-price auction, then there exists some valuation type \(v_i\) who has positive ex-post regret in some ex-post event. Since bidding equal to one’s valuation is a dominant strategy, this player can obtain a positive gain by switching her bid to \(v_i\). Let \(\sigma'\) be the new state after this player has switched to \(v_i\), *ceteris paribus*. Since the player gains a positive amount from this adjustment, the weight of the edge from \(\sigma\) to \(\sigma'\) in the weighted graph \(W_{SPA}\) is strictly less than \(\frac{1}{q}\). Now, suppose there exists a weighted cycle containing this outgoing edge. Then there must be an edge on this cycle such that our player changes her bid from \(v_i\). However, bidding \(v_i\) is a dominant strategy for the player of valuation \(v_i\). Therefore, the player can never gain a positive amount by switching her bid from \(v_i\). Hence, there cannot exist such a weighted cycle. As a result, \(\sigma\) passes the weighted cycle test and so it is not stochastically stable.

In a pure ex-post equilibrium, ex-post regrets of all types of all players are zero. Hence, any pure ex-post equilibrium is also a minimax regret equilibrium – but not vice versa. Thus in the second-price auction, every stochastically stable state is a minimax regret equilibrium.

**First-Price Auction:** Let \(M_{RM/FPA}(q, \alpha)\) be the Markov process with mistakes under random matching when the game is first-price auction. We have the following
Proposition 5.4. In the random-matching model, when the game is first-price auction, if \((\sigma_1, \ldots, \sigma_N)\) is stochastically stable in \(M_{RM/FPA}(q, \alpha)\) as \(q \to 0\), then \(\forall i \in N\) and \(v_i \in Z\), we have \(\sigma_i(v_i, a_i) > 0\) only if \(a_i \in \left[\frac{v_i - \delta}{2}, \frac{v_i + \delta}{2}\right]\cap Z\).

That is, when switches are unlikely events, all players will bid approximately half their valuations almost all the time. The intuition is as follows. Suppose in a state \(\sigma\), the player \(i\) of valuation \(v_i > \delta\) is bidding \(z < v_i\) (the proof takes care of all cases). The player gains from adjusting her bid upwards only when it converts her from a loser to a winner, while she gains from adjusting her bid downwards only when it reduces the price of winning the object. If \(z > \frac{v_i + \delta}{2}\), then the maximum gain from adjusting her bid downwards is equal to the highest price fall \(z - \delta\) (when the highest opposing bid encountered by her is 0). Since all 0 valuation types must bid 0 – the unique no-regret bid for 0 valuation players – there is a positive probability that the last highest bid encountered by the player is 0, and thus she could obtain the maximum gain of \(z - \delta\) by adjusting her bid downwards. Let \(\sigma'\) be the new state after this player reduces her bid to \(\delta\), ceteris paribus. Since the player gains \(z - \delta\) from this adjustment, the weight of the edge from \(\sigma\) to \(\sigma'\) in the weighted graph \(W_{FPA}\) is at most \(\frac{1}{z - \delta} < \frac{1}{\gamma}\). Now, consider any weighted cycle containing this outgoing edge. There must be an edge in this cycle such that our player increases her bid from some \(\tilde{z} < z\) to some \(\hat{z} \geq z\). However, the maximum gain of the player in this transition is \(v_i - \hat{z} \leq v_i - z < z - \delta\), when she turns from losing at \(\tilde{z}\) to winning at \(\hat{z}\). Hence, the weight of this edge in the weighted cycle is greater than the weight of the edge from \(\sigma\) to \(\sigma'\). Therefore, \(\sigma\) passes the weighted cycle test, and hence it is not stochastically stable. On the other hand, if \(z < \frac{v_i - \delta}{2}\), then the maximum gain from adjusting her bid upwards is at least \(v_i - z^*\), where \(z^*\) is the lowest bid that is at least \(\frac{v_i - \delta}{2}\). Now if there is another player \(j\) who bids \(z'\) such that \(z \leq z' < \frac{v_i - \delta}{2}\) – existence is shown in the proof – then there is a positive probability that these two players meet and player \(i\) loses the auction (this happens when all other players in this match have 0 valuations). Then player \(i\) can obtain the gain of \(v_i - z^*\) by adjusting her bid upwards to \(z^*\). Let \(\sigma'\) be the new state after this player increases her bid to \(z^*\), ceteris paribus. The weight of the edge from \(\sigma\) to \(\sigma'\) in the weighted graph \(W_{FPA}\) is at most \(\frac{1}{v_i - z^*} < \frac{1}{\gamma}\). Now, consider any weighted cycle containing this outgoing edge. There must be an edge in this cycle such that our player changes her bid from \(z^*\) to some \(\tilde{z} \neq z^*\). If \(\tilde{z} > z^*\), then the maximum gain of the player in this transition is \(v_i - \tilde{z} < v_i - z^*\); whereas if \(\hat{z} < z^*\), then the maximum gain of the player
is $z^* - \delta < v_i - z^*$ (since $z^* < \frac{u_i + \delta}{2}$). Hence, the weight of this edge in the weighted cycle is greater than the weight of the edge from $\sigma$ to $\sigma'$. Therefore, $\sigma$ passes the weighted cycle test, and hence it is not stochastically stable.

In fact, the stochastically stable bids (approximately) satisfy the following balance condition: at these bids, a player’s maximum gain from increasing her bid (approx. $v_i - a_i$) equals her maximum gain from decreasing it (approx. $a_i$), leading to $a_i = v_i/2$. Equivalently, at a stochastically stable bid, the maximum regret from not having used a higher bid is equal to the maximum regret from not having used a lower bid.

In the one-shot first-price auction under incomplete information with the set of valuations for each player equal to the unit interval $[0,1]$, the strategy profile in which each player bids half of her valuation is a minimax regret equilibrium. Hence, Proposition 5.4 predicts that if the players repeatedly face a random first-price auction environment and they are very slow to adapt their bids, then in the long run, the distributions of bids in the populations of players will approximately be a minimax regret equilibrium.

Bertrand Duopoly: Consider now the Bertrand duopoly game. If player $i$ with marginal cost $c_i$ were a monopolist, then the price of $\hat{p}(c_i) = p^*(c_i) - \frac{x - c_i}{2\sqrt{2}}$ gives her a payoff equal to half of her maximum payoff – it is the unique such price that is less than $p^*(c_i)$. Clearly, there does not exist any price in $P$ that equals $\hat{p}(c_i)$. Let $p^+(c_i)$ be the least element in $P$ that is at least $\hat{p}(c_i)$.

Proposition 5.5. In the random-matching model, when the game is Bertrand duopoly, if $(\sigma_1, \sigma_2)$ is stochastically stable in $\mathcal{M}_{\text{RM/BD}}(q, \alpha)$ as $q \to 0$, then $\forall i \in \{1, 2\}$ and $c_i \in C$, we have $\sigma_i(c_i, p_i) > 0$ only if $p_i \in [p^+(c_i), p^+(c_i) + \delta] \cap P$.

Thus, when switches are unlikely events, any player with marginal cost $c_i$ will post approximately the price $p^+(c_i)$ almost all the time. The intuition is as follows. Suppose in a state $\sigma$, player $i$ with marginal cost $c_i < x$ posts a price $p < p^+(c_i)$. Then $p < \hat{p}(c_i)$ and so the maximum gain from adjusting her price upwards is greater than half of her maximum monopoly payoff $\frac{(x - c_i)^2}{4y}$ (when the opponent’s price is greater than $p^*(c_i)$ and she switches from $p$ to $p^*(c_i)$). Since all players with marginal costs equal to $x$ must post the price $x$ – the unique no-regret price for such players –, there is a positive probability that the last price encountered by the player is $x$, and thus she could obtain the maximum gain by adjusting her price upwards to $p^*(c_i)$. Let $\sigma'$ be the new state after this player increases her price to $p^*(c_i)$, ceteris paribus. Since the player gains an amount greater than $\frac{(x - c_i)^2}{8y}$ from this adjustment, the weight of the edge from $\sigma$ to $\sigma'$ in the weighted graph $W_{\text{BD}}$ is less than $\frac{8y}{(x - c_i)^2} < \frac{1}{\gamma}$. Now,
consider any weighted cycle containing this outgoing edge. There must be an edge in this cycle such that our player decreases her price from some $\tilde{p} > p$ to some $\hat{p} \leq p$. However, the maximum gain of the player in this transition is less than $\frac{(x-c_i)^2}{8y}$ since $\hat{p} < \hat{p}(c_i)$. Hence, the weight of this edge in the weighted cycle is greater than the weight of the edge from $\sigma$ to $\sigma'$. Therefore, $\sigma$ passes the weighted cycle test, and hence it is not stochastically stable. On the other hand, if $p > p^+(c_i) + \delta$, then the maximum gain from adjusting her price downwards is greater than $\frac{(x-c_i)^2}{8y}$ (e.g., when she switches from zero market share at price $p$ to full market share at price $p^+(c_i)$). Now if there is a player $j$ who posts a price $p_j$ such that $p \geq p_j > p^+(c_i)$ – existence is shown in the proof that is available upon request –, then there is a positive probability that these two players meet and player $i$ obtains a zero market share. Then player $i$ can obtain a gain greater than $\frac{(x-c_i)^2}{8y}$ by adjusting her price downwards to $p^+(c_i)$. Let $\sigma'$ be the new state after this player decreases her price to $p^+(c_i)$, ceteris paribus. The weight of the edge from $\sigma$ to $\sigma'$ in the weighted graph $W_{BD}$ is less than $\frac{8y}{(x-c_i)^2} < \frac{1}{7}$. Now, consider any weighted cycle containing this outgoing edge. There must be an edge in this cycle such that our player changes her price from $p^+(c_i)$ to some $\tilde{p} \neq p^+(c_i)$. Thus either $p > p^+(c_i)$ or $p < \hat{p}(c_i)$. Hence, irrespective of the value of $p$, the maximum gain of the player is less than $\frac{(x-c_i)^2}{8y}$. Therefore, the weight of this edge in the weighted cycle is greater than the weight of the edge from $\sigma$ to $\sigma'$. So $\sigma$ passes the weighted cycle test, and hence it is not stochastically stable.

Like in the first-price auction, the stochastically stable prices in Bertrand competition (approximately) satisfy a balance condition: at these prices, a player’s maximum gain from increasing her price equals her maximum gain from decreasing it.

In the one-shot Bertrand duopoly game under incomplete information with the set of constant marginal costs equal to the interval $[0, x]$ (e.g., see Spulber, 1995), the strategy profile in which each player posts the price equal to $\hat{p}(c_i)$ is a minimax regret equilibrium. Hence, as in the second-price and first price auctions, Proposition 5.5 predicts that if the players repeatedly face a random Bertrand duopoly environment and they are very slow to adapt their prices, then in the long run, the distributions of prices in the populations of players will approximately be a minimax regret equilibrium.

\[ \diamond \]

$N \times 2$ Games: Let $G$ be a game with $N$ positions and two actions for each position, i.e., $|A_i| = 2, \forall i \in N$. We do not impose any restriction on the number of types of each player. The game $G$ satisfies strict dominance if for all $i \in N$, $a_i \in A_i$, there exists a type $t_i \in T_i$ such that $a_i$ is a strictly dominant action for type $t_i$. We have
the following result for $N \times 2$ games that satisfy strict dominance:

**Proposition 5.6.** Suppose $G$ is an $N \times 2$ game that satisfies strict dominance. If $\sigma \in \Sigma$ is stochastically stable, then $\sigma$ is a minimax regret equilibrium.\footnote{This result generalizes Corollary 3.3 in Saran and Serrano (2014) which proves the same result for the case of $2 \times 2$ games.}

### 6 Conclusion

We have studied a heuristic rule based on ex-post regret to grasp how to play games of incomplete information under private values. The conclusions depend on whether players interact within a fixed set (fixed matching) or they are randomly matched to play the game (random matching). When players do not make mistakes, the relevant long run predictions are minimal sets that are “closed under same or better reply” operations. For certain games, depending on the matching protocol and whether players make mistakes, the prediction boils down to pure Nash equilibrium, pure ex-post equilibrium or minimax regret equilibrium play. These three paradigms exhibit nice robustness properties in the sense that they are independent of beliefs about the exogenous uncertainty of type spaces. These predictions ought to be compared to the more standard one given by Bayesian equilibria. One obvious next step for further research would be to generalize these results for the case of interdependent values.

### 7 Proofs

**Proof of Proposition 3.2:** We first argue that $\hat{A}$ is a CUSOBR set if and only if whenever the dynamics $\mathcal{M}_{FM/G}(q)$ reaches any state in $\hat{A}$, it does not leave $\hat{A}$. Suppose $\hat{A}$ is such that whenever the dynamics reaches any state in $\hat{A}$, it never leaves $\hat{A}$. We argue that $R_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N) \subseteq \hat{A}$ for all $(a_1, \ldots, a_N) \in \hat{A}$. Pick any $(a'_1, \ldots, a'_N) \in R_{G(t_1, \ldots, t_N)}(a_1, \ldots, a_N)$. The argument is trivial if $(a'_1, \ldots, a'_N) = (a_1, \ldots, a_N)$. So suppose $(a'_1, \ldots, a'_N) \neq (a_1, \ldots, a_N)$. Let $I$ be the set of players such that $a'_i \neq a_i$. Then for all $i \in I$, action $a'_i$ is a better reply for type $t_i$ to $a_{-i}$ than $a_i$. There is a positive and independent probability that a player either does not switch her action or switches her action if she has positive regret. Hence, starting from state $(a_1, \ldots, a_N)$, there is a positive probability that, for all $i \in I$, the player of type $t_i$ switches her action to $a'_i$ while for all $j \notin I$, the player of type $t_j$ does not
switch her action. Hence, the dynamics moves from \((a_1, \ldots, a_N)\) to \((a'_1, \ldots, a'_N)\) in one period with a positive probability. Therefore, \((a'_1, \ldots, a'_N) \in \hat{A}\). Hence, \(\hat{A}\) is a CUSOBR set. Next, suppose \(\hat{A}\) is a CUSOBR set. Then for all \((a_1, \ldots, a_N) \in \hat{A}\), we have \(R_{\hat{A}}(a_1, \ldots, a_N) \subseteq \hat{A}\). Since a player switches her action with a positive probability only if she has positive regret, this means that once the dynamics reaches a state in \(\hat{A}\), it never leaves \(\hat{A}\). It follows that \(\hat{A}\) is a recurrent class if and only if it is a minimal CUSOBR set. 

**Proof of Lemma 3.4**: Suppose \(a^1 = (a^1_1, \ldots, a^1_N)\) is not a pure Nash equilibrium of the one-shot second-price auction in which the valuations \(\{v_1, \ldots, v_N\}\) of the players are common knowledge. Then there exists player \(i(1)\) who could gain by unilaterally deviating. Since bidding equal to one’s valuation is a weakly dominant strategy, unilaterally deviating to bidding \(v_i(1)\) is a better reply to \(a^1_{-i(1)}\) for player \(i(1)\). Let \(a^2\) be the bid profile such that player \(i(1)\) bids equal to her valuation and all other players \(j \neq i(1)\) continue to bid \(a^2_j = a^1_j\). If \(a^2\) is a pure Nash equilibrium of the game, then we are done. Otherwise, there exists a player \(i(2) \neq i(1)\) who could gain by unilaterally deviating when the bid profile is \(a^2\) (player \(i(1)\) cannot gain a positive amount by unilateral deviation since she is playing her weakly dominant strategy). Again, have \(i(2)\) bid truthfully and repeat the process. Since all players bidding truthfully is a pure Nash equilibrium of the game, it should be clear that in a finite number of steps, we will reach a bid profile \(a^L\) that is a pure Nash equilibrium of the game. Thus, the game is weakly acyclic under better replies. 

**Proof of Lemma 3.6**: Suppose \(a^1 = (a^1_1, \ldots, a^1_N)\) is not a pure Nash equilibrium of the one-shot first-price auction in which the valuations \(\{v_1, \ldots, v_N\}\) are common knowledge. Then there exists a player \(i(1)\) who could gain by unilaterally deviating. If player \(i(1)\) is getting a negative expected payoff at \(a^1\), let \(a^2_{i(1)} = 0\), which, with no other change to \(a^1\), leads to \(a^2\). If \(a^2\) is a Nash equilibrium, we are done, and if not, first make the same change in the bids of those players who are receiving a negative expected payoff. Suppose that, after such changes, the bid profile is \(a^l\). At \(a^l\) all players are receiving nonnegative payoffs. If \(a^l\) is a Nash equilibrium, we are done. If not, player \(i(l)\) has a profitable deviation, and there are only two possibilities: either she could gain by increasing her bid or decreasing her bid. To gain by increasing her bid, she must weakly outbid the highest bid, while to gain from decreasing her bid, either she must reduce her loss or increase her profit.

First, suppose player \(i(l)\) could gain by increasing her bid. Let \(a^{l*}\) be the highest

28
bid in the bid profile $a^l$. Thus, player $i(l)$ can weakly outbid the highest bid $a^i$ and make a positive profit. Then, in particular, bidding either $a^i$ or $a^i + \delta$ is a best reply to $a^l_{-i(l)}$ for player $i(l)$. Let $a^{l+1}$ be the bid profile such that player $i(l)$ bids such a best reply to $a^l_{-i(l)}$, and all other players $j \neq i(l)$ continue to bid $a^j_{l+1} = a^j_l$. Note that the result of this is that the price in the auction is at least as high as before. If $a^{l+1}$ is a pure Nash equilibrium of the game, then we are done. Otherwise, there exists a player $i(l + 1) \neq i(l)$ who could gain by unilaterally deviating when the bid profile is $a^{l+1}$. This means that player $i(l + 1)$ would like to weakly outbid the highest bid $a^{l+1}$ in the bid profile $a^{l+1}$. Let $a^{l+2}$ be a best reply to $a^{l+1}_{-i(l+1)}$, which leads to profile $a^{l+2}$. Note that, again, the price in the auction when the profile $a^{l+2}$ is used is at least as high as the previous price. Again, either $a^{l+2}$ is a Nash equilibrium and we are done, or it is not. But then, repeating the same argument a finite number of times, this process, based on a non-decreasing price sequence, must end at a Nash equilibrium.

Second, suppose player $i(l)$ can gain by decreasing her bid, and recall once again that all players are receiving nonnegative expected payoffs. If at $a^l$, player $i(l)$ can profitably deviate by decreasing her bid, noting that her expected payoff is nonnegative, she must be winning the auction but she can still win at a lower price. Thus, it must be that $a^i_{-i(l)} < a^i_{l(l)} \leq v_{i(l)}$, where $a^i_{-i(l)}$ is the highest bid among the bids $a^i_{-i(l)}$. Let $a^{l+1}_{i(l)} \geq a^i_{-i(l)}$ be a best reply to $a^l_{-i(l)}$, thus defining $a^{l+1}$. Now, either $a^{l+1}$ is a Nash equilibrium and we are done, or not. If not, it must be that one of the other players ($j \neq i(l)$) has a profitable deviation. If this deviation consists of increasing her bid, then we are back in the first case. So suppose that in $a^{l+1}$, no player has an improving deviation that consists of increasing her bid. Thus, for player $j \neq i(l)$, the improving deviation consists of decreasing her bid. Since player $i(l)$ bids the highest in $a^{l+1}$, it must be that $a^{l+1}_j = a^{l+1}_{i(l)}$ and player $j$ obtains a negative expected payoff in $a^{l+1}$. For any such $j$, bidding $a^{l+2}_j = a^{l+1}_{i(l)} - \delta$ is a better reply than $a^{l+1}_j$. In this way, change the bid of any such player $j$ to obtain the profile $a^{l+2}$. The highest (but not necessarily unique) player in $a^{l+2}$ is $i(l)$ and she does not have an improving deviation in $a^{l+2}$ (in particular, she cannot gain by reducing her bid to $a^{l+2}_{i(l)} - \delta$ because such a deviation at best ties her at the bid of $a^{l+2}_{i(l)} - \delta$). Clearly, none of the players who switched between $a^{l+1}$ and $a^{l+2}$ have an improving deviation. Finally, the rest of the players did not want to change their bid when the highest opposing bid was $a^{l+1}_{i(l)} = a^{l+2}_{i(l)}$ and so, they also do not have an improving deviation. Hence, $a^{l+2}$ is a Nash equilibrium.
Proof of Lemma 4.3: It is easy to see that if \((\sigma_1, \ldots, \sigma_N)\) is a pure ex-post equilibrium, then it must be that \(R_G(\sigma_1, \ldots, \sigma_N) = \{(\sigma_1, \ldots, \sigma_N)\}\). Now, suppose \((\sigma_1, \ldots, \sigma_N)\) is not a pure ex-post equilibrium. Then it must be that there exists \(i, (t_i, t_{-i}), (a_i, a_{-i}) \in \prod_{j \in N} A_j(\sigma_j(t_j))\) and \(a'_i \in A_i\) such that \(\pi_i(t_i, a'_i, a_{-i}) > \pi_i(t_i, a_i, a_{-i})\). Let \(\sigma'_i\) be obtained from \(\sigma_i\) when exactly one player of type \(t_i\) switches her action from \(a_i\) to \(a'_i\), ceteris paribus. Then \((\sigma'_i, \sigma_{-i}) \in R_G(\sigma_i, \sigma_{-i})\).

Proof of Proposition 4.4: We first argue that \(\hat{\Sigma}\) is a CUSOAEBR set if and only if whenever the dynamics \(M_{RM/G}(\eta)\) reaches any state in \(\hat{\Sigma}\), it does not leave \(\hat{\Sigma}\). Suppose \(\hat{\Sigma}\) is such that whenever the dynamics reaches any state in \(\hat{\Sigma}\), it never leaves \(\hat{\Sigma}\). We argue that \(R_G(\sigma_1, \ldots, \sigma_N) \subseteq \hat{\Sigma}, \forall (\sigma_1, \ldots, \sigma_N) \in \hat{\Sigma}\). Pick any \((\sigma'_1, \ldots, \sigma'_N) \in R_G(\sigma_1, \ldots, \sigma_N)\). If \((\sigma'_1, \ldots, \sigma'_N) = (\sigma_1, \ldots, \sigma_N)\), then the argument is trivial. So suppose \((\sigma'_1, \ldots, \sigma'_N) \neq (\sigma_1, \ldots, \sigma_N)\) and let \(I\) be the set of positions \(i\) such that \(\sigma'_i \neq \sigma_i\). We know that \((\sigma'_1, \ldots, \sigma'_N) \in \prod_{i \in N} \Sigma_i(\sigma_i)\) and that there exist \((t_1, \ldots, t_N)\) and \((a_1, \ldots, a_N) \in \prod_{i \in N} A_i(\sigma_i(t_i))\) such that for all \(i \in I\), we have \(t_i = t'_i\), \(a_i = a'_i\) and \(\pi_i(t'_i, a'_i, a_{-i}) > \pi_i(t_i, a_i, a_{-i})\).

In state \((\sigma_1, \ldots, \sigma_N)\), there is a positive probability that the type profile of the randomly matched players is \((t_1, \ldots, t_N)\) and these players play the action profile \((a_1, \ldots, a_N)\) since \((a_1, \ldots, a_N) \in \prod_{i \in N} A_i(\sigma_i(t_i))\). There is a positive and independent probability that a player either does not switch her action or switches her action if she has positive regret. Hence, there is a positive probability that, for all \(i \in I\), the player of type \(t_i\) switches to \(a_i^{2*}\) while for all \(j \notin I\), the player of type \(t_j\) continues to play \(a_j\). Furthermore, there is no change in the distributions of actions in all other populations because none of the players from these populations is matched. Hence, there is a positive probability that for all \(i \in I\), the distributions of actions in population \(\mathcal{P}_i\) in the next period will be \(\sigma'_i\) while for all \(j \notin I\), the distributions of actions in population \(\mathcal{P}_j\) in the next period will remain \(\sigma_j\). Thus, the dynamics will move from \((\sigma_1, \ldots, \sigma_N)\) to \((\sigma'_1, \ldots, \sigma'_N)\) in one period with a positive probability. Therefore, it must be that \((\sigma'_1, \ldots, \sigma'_N) \in \hat{\Sigma}\). Hence, \(\hat{\Sigma}\) is a CUSOAEBR set. Next, suppose \(\hat{\Sigma}\) is a CUSOAEBR set. Then for all \((\sigma_1, \ldots, \sigma_N) \in \hat{\Sigma}\), we have \(R_G(\sigma_1, \ldots, \sigma_N) \subseteq \hat{\Sigma}\). A player switches her action with a positive probability only if she has positive regret and at most a single player in any population \(\mathcal{P}_i\) switches her action in one period (since only one such player is matched). Hence, once the dynamics reaches a state in \(\hat{\Sigma}\), it never leaves \(\hat{\Sigma}\). It follows that \(\hat{\Sigma}\) is a recurrent class if and only if it is a minimal CUSOAEBR set.
Proof of Lemma 4.5: Suppose $\hat{\Sigma}$ is a minimal CUSOAEBR set of the second-price auction but it is no singleton. Pick any $\sigma^1 = (\sigma^1_1, \ldots, \sigma^1_N) \in \hat{\Sigma}$. Clearly, $\sigma^1$ is not a pure ex-post equilibrium of the second-price auction. Then there exists $i$, an ex-post realization of types $(v_i, v_{-i})$, and action profile $(a_i, a_{-i}) \in \prod_{j \in N} A_j(\sigma^1_j(v_j))$, such that player $i$ could gain by unilaterally deviating in this ex-post event. Since bidding equal to one’s valuation is a weakly dominant strategy, it must be that $a_i \neq v_i$. Moreover, unilaterally deviating to bidding $v_i$ is a better reply for player $i$ in this ex-post event. Let $\sigma^2$ be the strategy profile such that $\sigma^2_j = \sigma^1_j$ for all $j \neq i$ and $\sigma^2_i$ is adjacent to $\sigma^1_i$ with $\sigma^2_i(v_i, a_i) < \sigma^1_i(v_i, a_i)$ and $\sigma^2_i(v_i, v_i) > \sigma^1_i(v_i, v_i)$. Then $\sigma^2 \in R_{SPA}(\sigma^1)$, and hence $\sigma^2 \in \hat{\Sigma}$. If $\sigma^2$ is a pure ex-post equilibrium of the game, then we will obtain a contradiction. If not, then repeat the above argument. Since the strategy profile in which all types of all players only bid truthfully is a pure ex-post equilibrium of the game, it should be clear that in a finite number of steps, we will reach a strategy profile $\sigma^L$ that is a pure ex-post equilibrium of the game. But $\sigma^L \in \hat{\Sigma}$, which is a contradiction. Thus, $\hat{\Sigma}$ must be singleton. \hfill $\square$

Proof of Lemma 4.7: Recall the definition of $\bar{R}_G$ and accordingly define $\bar{R}_{FPA}$. Let $\hat{\sigma} = (\hat{\sigma}_1, \ldots, \hat{\sigma}_N)$ be such that $\hat{\sigma}_i(0, 0) = \sigma_i(\delta, 0) = 1$ and $\hat{\sigma}_i(v_i, \delta) = 1, \forall v_i > \delta$. Pick any $\sigma = (\sigma_1, \ldots, \sigma_N)$ and consider the iteration

$$\bar{R}_{FPA}({\sigma}) \subseteq \bar{R}_{FPA}^2({\sigma}) \subseteq \ldots \subseteq \bar{R}_{FPA}^{l}({\sigma}) \subseteq \ldots$$

We argue that there exists $l$ such that $\hat{\sigma} \in \bar{R}_{FPA}^{l}({\sigma})$. Suppose $\sigma$ is such that there exists a 0 valuation type of some player who plays a positive bid with a positive probability. Let $z$ be the greatest positive bid played by any 0 valuation type of any player. Without loss of generality, let player $i$ be such that $\sigma_i(0, z) > 0$. Let $\sigma^1_i \in \Sigma_i(\sigma_i)$ be such that $\sigma^1_i(0, 0) > \sigma_i(0, 0) \text{ and } \sigma^1_i(0, z) < \sigma_i(0, z)$. Then the strategy profile $\sigma^1 = (\sigma^1_1, \ldots, \sigma^1_N)$ such that $\sigma^1_j = \sigma_j, \forall j \neq i$ is in $\bar{R}_{FPA}({\sigma})$. This is because in the ex-post event $(v_1, \ldots, v_N)$ such that $v_j = 0, \forall j \in N$ and $(a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\sigma^1_j(0))$ such that $a_i = z$, player $i$ obtains a negative expected payoff, and hence the bid of 0 is a better reply to $a_{-i}$. If $\sigma^1$ is such that there exists a 0 valuation type of some player who plays a positive bid with a positive probability, then repeat this argument until we reach $\sigma^l$ such that all 0 valuation types of all players surely bid 0. By construction, $\sigma^l \in \bar{R}_{FPA}^{l}({\sigma})$. Suppose $\sigma^l$ is such that there exists a $\delta$ valuation type of some player who plays a positive bid with a positive probability. Let $z'$ be the greatest positive bid played by any $\delta$ valuation type of any player.
Without loss of generality, let player \( i \) be such that \( \sigma_i(\delta, z') > 0 \). Let \( \sigma_i^{l+1} \in \Sigma_i(\sigma_i^l) \) be such that \( \sigma_i^{l+1}(\delta, 0) > \sigma_i^l(\delta, 0) \) and \( \sigma_i^{l+1}(\delta, z') < \sigma_i^l(\delta, z') \). Then the strategy profile \( \sigma^{l+1} = (\sigma_1^{l+1}, \ldots, \sigma_N^{l+1}) \) such that \( \sigma_j^{l+1} = \sigma_j^l, \forall j \neq i \) is in \( \tilde{R}_{FPA}(\{\sigma^l\}) \), and hence in \( \tilde{R}_{FPA}([\sigma]) \). This is because in the ex-post event \( (v_1, \ldots, v_N) \) such that \( v_j = 0, \forall j \neq i \) and \( v_i = \delta \), and \( (a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\sigma_j^l(v_j)) \) such that \( a_j = 0, \forall j \neq i \) and \( a_i = z' \), player \( i \) obtains at most an expected payoff of 0, and hence the bid of 0 is a better reply to \( a_{\sim i} \). If \( \sigma^{l+1} \) is such that there exists a \( \delta \) valuation type of some player who plays a positive bid with a positive probability, then repeat this argument until we reach \( \sigma^{l''} \) such that all 0 and \( \delta \) valuation types of all players surely bid 0. By construction, \( \sigma^l \in \tilde{R}_{FPA}(\{\sigma\}) \). Suppose \( \sigma^l \) is such that there exists a valuation type \( v_i' > 2\delta \) of player \( i \) who bids \( z'' \neq \delta \) with a positive probability. Let \( \sigma_i^{l+1} \in \Sigma_i(\sigma_i^l) \) be such that \( \sigma_i^{l+1}(v_i', \delta) > \sigma_i^l(v_i', \delta) \) and \( \sigma_i^{l+1}(v_i', z'') < \sigma_i^l(v_i', z'') \). Then the strategy profile \( \sigma^{l+1} = (\sigma_1^{l+1}, \ldots, \sigma_N^{l+1}) \) such that \( \sigma_j^{l+1} = \sigma_j^l, \forall j \neq i \) is in \( \tilde{R}_{FPA}(\{\sigma^l\}) \), and hence in \( \tilde{R}_{FPA}([\sigma]) \). This is because in the ex-post event \( (v_1, \ldots, v_N) \) such that \( v_j = 0, \forall j \neq i \) and \( v_i = v_i' \), and \( (a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\sigma_j^l(v_j)) \) such that \( a_j = 0, \forall j \neq i \) and \( a_i = z'' \), the expected payoff of player \( i \) is either \( v_i'/N \) if \( z'' = 0 \) or \( v_i' - z'' \) if \( z'' > \delta \), and hence, in any case, the bid of \( \delta \) with the payoff \( v_i' - \delta \) is a better reply to \( a_{\sim i} \). If \( \sigma^{l+1} \) is such that there exists a valuation type \( v_i' > 2\delta \) of some player \( j \) who plays a bid other than \( \delta \) with a positive probability, then repeat this argument until we reach \( \sigma^{l''} \) such that all 0 and \( \delta \) valuation types of all players surely bid 0 and all valuation types of all players with valuations greater than \( 2\delta \) surely bid \( \delta \). By construction, \( \sigma^{l''} \in \tilde{R}_{FPA}(\{\sigma\}) \). Suppose \( \sigma^{l''} \) is such that there exists a valuation type \( 2\delta \) of player \( i \) who plays a bid \( \hat{z} \neq \delta \) with a positive probability. Let \( \sigma_i^{l''+1} \in \Sigma_i(\sigma_i^{l''}) \) be such that \( \sigma_i^{l''+1}(2\delta, \delta) > \sigma_i^{l''}(2\delta, \delta) \) and \( \sigma_i^{l''+1}(2\delta, \hat{z}) < \sigma_i^{l''}(2\delta, \hat{z}) \). Then the strategy profile \( \sigma^{l''+1} = (\sigma_1^{l''+1}, \ldots, \sigma_N^{l''+1}) \) such that \( \sigma_j^{l''+1} = \sigma_j^{l''}, \forall j \neq i \) is in \( \tilde{R}_{FPA}(\{\sigma^{l''}\}) \), and hence in \( \tilde{R}_{FPA}([\sigma]) \). This is because in the ex-post event \( (v_1, \ldots, v_N) \) such that \( v_j = 3\delta, \forall j \neq i \) and \( v_i = 2\delta \), and \( (a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\sigma_j^{l''}(v_j)) \) such that \( a_j = \delta, \forall j \neq i \) and \( a_i = \hat{z} \), the payoff of player \( i \) is at most 0, and hence the bid of \( \delta \) is a better reply to \( a_{\sim i} \). If \( \sigma^{l''+1} \) is such that there exists valuation type \( 2\delta \) of some player \( j \) who plays a bid other than \( \delta \) with a positive probability, then repeat this argument until we reach \( \sigma \). By construction, \( \sigma \in \tilde{R}_{FPA}(\{\sigma\}) \) for some \( \hat{l} \).

It follows that any CUSOAEBR set must contain \( \sigma \), and hence there exists a unique minimal CUSOAEBR set.

Pick any player \( i \) of valuation \( v_i' \in [3\delta, 1] \) and \( a_i \in [\delta, v_i' - 2\delta] \cap \mathbb{Z} \). If \( v_i' = 3\delta \), then \( a_i = \delta \) and we already have \( \hat{\sigma} \in \tilde{\Sigma} \) such that \( \hat{\sigma}_i(3\delta, \delta) = 1 \). So pick \( v_i' \in [4\delta, 1] \). Let \( \hat{\sigma}_i^1 \in \Sigma_i(\hat{\sigma}_i) \) be such that \( \hat{\sigma}_i^1(v_i', 2\delta) > \hat{\sigma}_i(v_i', 2\delta) \) and \( \hat{\sigma}_i^1(v_i', \delta) < \hat{\sigma}_i(v_i', \delta) \). Then the
strategy profile $\hat{\sigma}^1 = (\hat{\sigma}_1^1, \ldots, \hat{\sigma}_N^1)$ such that $\hat{\sigma}_j^1 = \hat{\sigma}_j, \forall j \neq i$ is in $\tilde{R}_{SPA}(\{\hat{\sigma}\})$. This is because in the ex-post event $(v_1, \ldots, v_N)$ such that $v_j = 3\delta, \forall j \neq i$ and $v_i = v_i'$, and $(a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\hat{\sigma}_j(v_j))$ such that $a_j = \delta, \forall j \in N$, player $i$ obtains an expected payoff of $(v_i' - \delta)/N$ and hence, the bid of $2\delta$ with a payoff of $v_i' - 2\delta$ is a better reply to $a_{-i}$. By repeating this argument we will reach $\hat{\sigma}^l$ such that for all $j \in N$, $\hat{\sigma}_j^l(v_j, 0) = 1, \forall v_j \in [0, \delta]$, $\hat{\sigma}_j^l(v_j, \delta) = 1, \forall v_j \in [2\delta, 3\delta]$ and $\hat{\sigma}_j^l(v_j, 2\delta) = 1, \forall v_j \in [4\delta, 1]$. By construction, $\hat{\sigma}^l \in \tilde{R}_{SPA}(\{\hat{\sigma}\})$. Now, pick $v_i' \in [5\delta, 1]$. Let $\tilde{\sigma}_i^{l+1} \in \Sigma_i(\hat{\sigma}_i^l)$ be such that $\tilde{\sigma}_i^{l+1}(v_i', 3\delta) > \tilde{\sigma}_i^l(v_i', 3\delta)$ and $\tilde{\sigma}_i^{l+1}(v_i', 2\delta) < \tilde{\sigma}_i^l(v_i', 2\delta)$. Then the strategy profile $\tilde{\sigma}^{l+1} = (\tilde{\sigma}_1^{l+1}, \ldots, \tilde{\sigma}_N^{l+1})$ such that $\tilde{\sigma}_j^{l+1} = \tilde{\sigma}_j, \forall j \neq i$ is in $\tilde{R}_{SPA}(\{\hat{\sigma}^l\})$, and hence in $\tilde{R}_{SPA}(\{\hat{\sigma}\})$. This is because in the ex-post event $(v_1, \ldots, v_N)$ such that $v_j = 4\delta, \forall j \neq i$ and $v_i = v_i'$, and $(a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\hat{\sigma}_j(v_j))$ such that $a_j = 2\delta, \forall j \in N$, player $i$ obtains an expected payoff of $(v_i' - 2\delta)/N$, and hence the bid of $3\delta$ with a payoff of $v_i' - 3\delta$ is a better reply to $a_{-i}$. By repeating this argument we will reach $\tilde{\sigma}^l$ such that for all $j \in N$, $\tilde{\sigma}_j^l(v_j, 0) = 1, \forall v_j \in [0, \delta]$, $\tilde{\sigma}_j^l(v_j, \delta) = 1, \forall v_j \in [2\delta, 3\delta]$, $\tilde{\sigma}_j^l(4\delta, 2\delta) = 1$, and $\tilde{\sigma}_j^l(v_j, 3\delta) = 1, \forall v_j \in [5\delta, 1]$. By construction, $\sigma^l \in \tilde{R}_{SPA}(\{\tilde{\sigma}\})$. It should be clear that the statement in the lemma can be proved after a finite number of repetitions of the above argument. 

Proof of Proposition 5.3: We use the weighted cycle test to prove this proposition. However, we need to modify the weighted graph $W_{SPA}$ to incorporate the fact that in the event of a tie at the winning bid, the realized payoff of a player need not equal her expected payoff (where the expectation is with respect to the tie-breaking lottery). This is achieved by redefining $R_{SPA}$ as follows: $\sigma' \in R_{SPA}(\sigma)$ if and only if $\sigma'_i \in \Sigma_i(\sigma_i), \forall i$, and if $\sigma' \neq \sigma$, then letting $I = \{i|\sigma_i \neq \sigma'_i\}$, there must exist $(v_1, \ldots, v_N)$ and $(a_1, \ldots, a_N) \in \prod_{i \in N} A_i(\sigma_i(v_i))$ such that for all $i \in I$, we have $v_i = v_i^*$ (where $v_i^*$ is the unique valuation such that $\sigma'_i(v_i^*) \neq \sigma_i(v_i^*)$), $a_i = a_i^{1*}$ and the expected payoff of player $i$ of valuation $v_i^*$ from bidding $a_i^{2*}$ against $a_{-i}$ is greater than her payoff after some realization of the tie-breaking lottery following her bid of $a_i^{1*}$ against $a_{-i}$. Now, use this redefined $R_{SPA}$ to modify the weighted graph $W_{SPA}$ by adding an edge from $\sigma$ to $\sigma' \neq \sigma$ if and only if $\sigma' \in R_{SPA}(\sigma)$. The weight of this edge is still the resistance $r(\sigma, \sigma')$. The weighted cycle test is also valid for this modified weighted graph, which we continue to denote by $W_{SPA}$.

Consider any state $\sigma$ that is not a pure ex-post equilibrium of the second-price auction. Then there exists $i, (v_i, v_{-i}), (a_i, a_{-i}) \in \prod_{j \in N} A_j(\sigma_j(v_j))$, and action $a'_i$ such that the expected payoff of valuation type $v_i$ from bidding $a'_i$ against $a_{-i}$ is greater than her expected payoff from bidding $a_i$ against $a_{-i}$. Hence, in some ex-post event
(i.e., after the realization of the tie-breaking lottery), the payoff obtained by \( v_i \) when she bids \( a_i \) against \( a_{-i} \) is lower than her expected payoff from bidding \( a'_i \) against \( a_{-i} \). Since bidding equal to one's valuation is a weakly dominant strategy, it must be that \( a_i \neq v_i \) and it is without loss of generality to take \( a'_i = v_i \). Let \( \sigma_i^1 \in \Sigma_i(\sigma_i) \) be such that \( \sigma_i^1(v_i, v_i) > \sigma_i(v_i, v_i) \) and \( \sigma_i^1(v_i, a_i) < \sigma_i(v_i, a_i) \). Then the strategy profile \( \sigma^1 = (\sigma_1^1, \ldots, \sigma_N^1) \) such that \( \sigma_i^1 = \sigma_j, \forall j \neq i \) is in \( R_{SPA}(\sigma) \). This is because in the ex-post event \((v_i, v_{-i})\) and \((a_i, a_{-i}) \in \prod_{j \in N} A_j(\sigma_j(v_j)) \), there is a positive probability that player \( i \) has a positive ex-post regret from bidding \( a_i \) instead of \( v_i \) against \( a_{-i} \).

So \( r(\sigma, \sigma^1) < \frac{1}{\gamma} \) since only a single player of valuation \( v_i \) in population \( \mathcal{P}_i \) switches from \( a_i \) to \( v_i \) and there exists an ex-post event in state \( \sigma \) in which this player has positive regret from bidding \( a_i \) instead of \( v_i \). Thus in \( W_{SPA} \), there exists an outgoing edge from \( \sigma \) to \( \sigma^1 \) with a weight of less than \( \frac{1}{\gamma} \). Now, consider any weighted cycle containing this outgoing edge. Since \( \sigma_i^1(v_i, v_i) > \sigma_i(v_i, v_i) \), it must be that there exist two states \( \sigma^i, \sigma^{i+1} \) such that \( \sigma^{i+1} \in R_{SPA}(\sigma^i) \) and \( \sigma^{i+1}_i(v_i, v_i) < \sigma_i^i(v_i, v_i) \). That is, there exists a player of valuation \( v_i \) in population \( \mathcal{P}_i \) who switches from bidding \( v_i \) to some \( a''_i \). However, bidding \( v_i \) is a weakly dominant strategy for player of valuation \( v_i \), which implies that in state \( \sigma^i \) there does not exist any ex-post event in which \( a''_i \) is a better reply than \( v_i \) for type \( v_i \). This contradicts the fact that \( \sigma^{i+1} \in R_{SPA}(\sigma^i) \).

Thus, the state \( \sigma \) passes the weighted cycle test, and hence it is not stochastically stable.

\[ \square \]

**Proof of Proposition 5.4:** First, redefine the weighted graph \( W_{FPA} \) as we did for the second-price auction in the proof of Proposition 5.3. We use the weighted cycle test on this modified weighted graph, which we continue to denote by \( W_{FPA} \).

We prove by induction that a state \( \sigma \) is stochastically stable only if \( \forall n = 0, \ldots, \frac{1}{\delta}, \) the support of \( \sigma_i(n\delta), A_i(\sigma_i(n\delta)) \subseteq \left[ \frac{(n-1)\delta}{2}, \frac{(n+1)\delta}{2} \right] \cap Z, \forall i \in N \).

First, the statement is true for \( n = 0 \). To see this, note that it is weakly dominant for a player of 0 valuation to bid 0, and if \( \sigma \) is such that some player of 0 valuation bids a positive amount, then the highest such bidder has a positive probability of having a positive regret from her bid in any ex-post event in which all matched players have 0 valuations. Now using an argument similar to the one in the proof of Proposition 5.3, we can easily argue that \( \sigma \) passes the weighted cycle test, and hence it is not stochastically stable.

Now, suppose the statement is true for all \( n' \leq n \). We argue that it is also true for \( n + 1 \). Consider a state \( \sigma \) such that \( \forall j \in N \) and \( v_j \leq n\delta \), we have \( A_j(\sigma_j(v_j)) \subseteq \left[ \frac{v_j-\delta}{2}, \frac{v_j+\delta}{2} \right] \cap Z \) but there exists a player of valuation \( (n+1)\delta \) in population \( \mathcal{P}_i \) who
bids \( z \notin \left[ \frac{n\delta}{2}, \frac{(n+2)\delta}{2} \right] \cap \mathbb{Z} \).

First suppose \( z > \frac{(n+2)\delta}{2} \). Let \( \sigma_i^1 \in \Sigma_i(\sigma_i) \) be such that \( \sigma_i^1((n+1)\delta, \delta) > \sigma_i((n+1)\delta, \delta) \) and \( \sigma_i^1((n+1)\delta, z) < \sigma_i((n+1)\delta, z) \). Then the strategy profile \( \sigma^1 = (\sigma_1^1, \ldots, \sigma_N^1) \) such that \( \sigma_j^1 = \sigma_j, \forall j \neq i \) is in \( R_{FPA}(\sigma) \). To see this, consider the ex-post event \((v_1, \ldots, v_N) \) such that \( v_j = 0, \forall j \neq i \) and \( v_i = (n+1)\delta \), and \( (a_1, \ldots, a_N) \in \prod_{j \in N} A_j(\sigma_i(v_j)) \) such that \( a_j = 0, \forall j \neq i \) and \( a_i = z \). Then bidding \( \delta \) is a better reply to \( a_i \) than \( z \) and player \( i \)'s ex-post regret is \( z - \delta \). Hence, \( r(\sigma, \sigma^1) \leq \frac{1}{z - \delta} < \frac{1}{\gamma} \) since only a single player of valuation \((n+1)\delta \) in population \( P_i \) switches from the bid of \( z \) to the bid of \( \delta \) and there exists an ex-post event in state \( \sigma \) in which this player has positive ex-post regret of \( z - \delta \) from bidding \( z \) instead of \( \delta \). Thus, in \( W_{FPA} \), there exists an outgoing edge from \( \sigma \) to \( \sigma^1 \) with a weight of less than \( \frac{1}{\gamma} \). Now, consider any weighted cycle containing this outgoing edge. Since \( \sum_{z' \geq z} \sigma_i^1((n+1)\delta, z') < \sum_{z' \geq z} \sigma_i((n+1)\delta, z') \), it must be that there exist two states \( \sigma^1, \sigma^{l+1} \) such that \( \sigma^{l+1} \in R_{FPA}(\sigma^l) \) and \( \sum_{z' \geq z} \sigma_i^{l+1}((n+1)\delta, z') > \sum_{z' \geq z} \sigma_i^l((n+1)\delta, z') \). That is, there exists a player of valuation \((n+1)\delta \) in population \( P_i \) who switches from bidding \( \tilde{z} < z \) to some \( \hat{z} > z \). However, the player \( i \) of valuation \((n+1)\delta \) can gain a positive amount from increasing her bid from \( \tilde{z} \) to \( \hat{z} \) against any \( a_i \) only if \( \hat{z} < (n+1)\delta \) and the maximum such gain equals \((n+1)\delta - \tilde{z} \) (when she turns from losing at \( \tilde{z} \) to winning at \( \hat{z} \)). Since \( \hat{z} > \frac{(n+2)\delta}{2} \), the maximum gain from such a switch is less than \( \frac{n\delta}{2} \). Hence, \( r(\sigma^l, \sigma^{l+1}) > \frac{2}{n\delta} > \frac{1}{z - \delta} \) since \( z > \frac{(n+2)\delta}{2} \). Thus, the state \( \sigma \) passes the weighted cycle test, and hence it is not stochastically stable.

Finally, suppose \( z < \frac{n\delta}{2} \). This is possible only if \( n \geq 1 \). Let \( z^* \) be the best bid greater than or equal to \( \frac{n\delta}{2} \). Let \( \sigma_i^1 \in \Sigma_i(\sigma_i) \) be such that \( \sigma_i^1((n+1)\delta, z^*) > \sigma_i((n+1)\delta, z^*) \) and \( \sigma_i^1((n+1)\delta, z) < \sigma_i((n+1)\delta, z) \).

**Case 1:** \( n \) is odd. Then \( z \leq \frac{(n-1)\delta}{2} \) and \( z^* = \frac{(n+1)\delta}{2} \). Pick a \( j \neq i \) and \( v_j = (n-1)\delta \). By the induction hypothesis, there exists a \( z' \in \left[ \frac{n\delta}{2}, \frac{n\delta}{2} \right] \cap \mathbb{Z} \) such that \( \sigma_j(v_j, z') > 0 \). Then \( z' = \frac{(n-1)\delta}{2} \). Consider the ex-post event \((v_1, \ldots, v_N) \) such that \( v_k = 0, \forall k \neq i, j \), \( v_j = (n-1)\delta \) and \( v_i = (n+1)\delta \), and \( (a_1, \ldots, a_N) \in \prod_{k \in N} A_k(\sigma_k(v_k)) \) such that \( a_k = 0, \forall k \neq i, j \), \( a_j = z' \) and \( a_i = z \). Since \( z \leq z' \), with a positive probability, player \( i \) loses the auction. Therefore, in this ex-post stage, bidding \( z^* \) is a better reply to \( a_i \) than \( z \) and player \( i \)'s ex-post regret is \((n+1)\delta - z^* \). Hence, letting \( \sigma^1 = (\sigma_1^1, \ldots, \sigma_N^1) \) such that \( \sigma_j^1 = \sigma_j, \forall j \neq i \), we have \( \sigma^1 \in R_{FPA}(\sigma) \) and \( r(\sigma, \sigma^1) \leq \frac{1}{(n+1)\delta - z^*} < \frac{1}{\gamma} \) since only a single player of valuation \((n+1)\delta \) in population \( P_i \) switches from the bid of \( z \) to the bid of \( z^* \) and there exists an ex-post event in state \( \sigma \) in which this player has positive ex-post regret of \((n+1)\delta - z^* \) from bidding \( z \) instead of \( z^* \).
Thus, in $W_{FPA}$, there exists an outgoing edge from $\sigma$ to $\sigma^1$ with a weight of less than $\frac{1}{\gamma}$. Now, consider any weighted cycle containing this outgoing edge. Since $\sigma^1_i((n+1)\delta, z^*) > \sigma_i((n+1)\delta, z^*)$, it must be that there exist two states $\sigma^t, \sigma^{t+1}$ such that $\sigma^{t+1} \in R_{FPA}(\sigma^t)$ and $\sigma^{t+1}_i((n+1)\delta, z^*) < \sigma^t_i((n+1)\delta, z^*)$. That is, there exists a player of valuation $(n+1)\delta$ in population $P_i$ who switches from bidding $z^*$ to some $\hat{z} \neq z^*$. Suppose $\hat{z} > z^*$. The player $i$ of valuation $(n+1)\delta$ can gain a positive amount from increasing her bid from $z^*$ to $\hat{z}$ against any $a_{-i}$ only if $\hat{z} < (n+1)\delta$ and the maximum such gain equals $(n+1)\delta - \hat{z} < (n+1)\delta - z^*$ (when she turns from losing at $z^*$ to winning at $\hat{z}$). Next, suppose $\hat{z} < z^*$. The player $i$ of valuation $(n+1)\delta$ can gain a positive amount from decreasing her bid from $z^*$ to $\hat{z}$ against any $a_{-i}$ only if $\hat{z} > (n+1)\delta$ and the maximum such gain equals $z^* - \delta < (n+1)\delta - z^*$ (when turns from winning at $z^*$ to winning at the bid of $\delta$).

Hence, $r(\sigma^t, \sigma^{t+1}) > \frac{1}{(n+1)\delta - z^*} \geq r(\sigma, \sigma^1)$. Thus, the state $\sigma$ passes the weighted cycle test, and so it is not stochastically stable.

**Case 2:** $n$ is even. Then $n \geq 2$, $z \leq \frac{(n-2)\delta}{2}$ and $z^* = \frac{n\delta}{2}$. Pick a $j \neq i$ and $v_j = (n-2)\delta$. By the induction hypothesis, there exists a $z' \in \left[\frac{(n-3)\delta}{2}, \frac{(n-1)\delta}{2}\right] \cap Z$ such that $\sigma_j(v_j, z') > 0$. Then $z' = \frac{(n-2)\delta}{2}$. Consider the ex-post event $(v_1, \ldots, v_N)$ such that $v_k = 0, \forall k \neq i, j, v_j = (n-2)\delta$ and $v_i = (n+1)\delta$, and $(a_1, \ldots, a_N) \in \prod_{k \in N} A_k(\sigma_k(v_k))$ such that $a_k = 0, \forall k \neq i, j, a_j = z'$ and $a_i = z$. Since $z \leq z'$, with a positive probability, player $i$ loses the auction. Therefore, in this ex-post stage, bidding $z^*$ is a better reply to $a_{-i}$ than $z$ and player $i$’s ex-post regret is $(n+1)\delta - z^*$.

Hence, letting $\sigma^1 = (\sigma^1_1, \ldots, \sigma^1_N)$ such that $\sigma^1_j = \sigma_j, \forall j \neq i$, we have $\sigma^1 \in R_{FPA}(\sigma)$ and $r(\sigma, \sigma^1) \leq \frac{1}{(n+1)\delta - z^*} < \frac{1}{\gamma}$. Thus, in $W_{FPA}$, there exists an outgoing edge from $\sigma$ to $\sigma^1$ with a weight of less than $\frac{1}{\gamma}$. Like in Case 1, we can easily argue that in any weighted cycle containing this outgoing edge, there exists an edge with weight greater than $r(\sigma, \sigma^1)$. Hence, $\sigma$ is not stochastically stable.

**Proof of Proposition 5.6:** Since $G$ is $N \times 2$ game, let $A_i = \{a_i, a'_i\}$. Consider any $\sigma$ such that $\sigma$ is stochastically stable.

We first argue that $\sigma$ is such that $\bigcup_{t_i \in T_i} A_i(\sigma_i(t_i)) = A_i$ for all $i$. Suppose not, i.e., without loss of generality, there exists an $i$ such that $\bigcup_{t_i \in T_i} A_i(\sigma_i(t_i)) = \{a_i\}$. Since $G$ satisfies strict dominance, there exists an $t'_i$ such that $a'_i$ is a strictly dominant action for type $t'_i$. Let $\sigma^1 = (\sigma^1_1, \sigma_{-i})$, where $\sigma^1_i \in \Sigma_i(\sigma_i)$ is such that $\sigma^1_i(t'_i, a'_i) > \sigma_i(t'_i, a'_i)$ and $\sigma^1_i(t'_i, a_i) < \sigma_i(t'_i, a_i)$. In state $\sigma$, consider the ex-post event $(t'_i, t_{-i})$ such that player $i$ plays $a_i$ and other players play some $a_{-i} \in \prod_{j \neq i} \left(\bigcup_{t_j \in T_j} A_j(\sigma_j(t_j))\right)$. Then playing $a'_i$ is a better reply to $a_{-i}$ than $a_i$ and player $i$’s ex-post regret is positive.
Hence, \( \sigma^1 \in R_G(\sigma) \) and \( r(\sigma, \sigma^1) < \frac{1}{\gamma} \) since only a single player of type \( t'_i \) in population \( \mathcal{P}_i \) switches her action from \( a_i \) to \( a'_i \) and there exists an ex-post event in state \( \sigma \) in which this player has positive ex-post regret from playing \( a_i \) instead of \( a'_i \). Thus, in the graph \( W_G \), there exists an outgoing edge from \( \sigma \) to \( \sigma^1 \) with a weight of less than \( \frac{1}{\gamma} \). Now, consider any weighted cycle containing this outgoing edge. Since \( \sigma^1(t'_i, a'_i) > \sigma_i(t'_i, a'_i) \), it must be that there exist two states \( \sigma^l, \sigma^{l+1} \) in this weighted cycle such that \( \sigma^{l+1} \in R_G(\sigma^l) \) and \( \sigma^{l+1}(t'_i, a'_i) < \sigma^l(t'_i, a'_i) \). That is, there exists a player of type \( t'_i \) in population \( \mathcal{P}_i \) who switches her action from \( a'_i \) to \( a_i \). However, \( a'_i \) is strictly dominant for \( t'_i \), which implies that in state \( \sigma^l \), there does not exist any ex-post event in which \( a_i \) is a better reply than \( a'_i \) for type \( t'_i \). This contradicts the fact that \( \sigma^{l+1} \in R_G(\sigma^l) \). Thus, the state \( \sigma \) passes the weighted cycle test, and hence it is not stochastically stable.

For all \( a_{-i} \), let \( \tilde{\Delta}_i(t_i, a_i, a'_i, a_{-i}) = \max\{\Delta_i(t_i, a_i, a'_i, a_{-i}), 0\} \) and \( \tilde{\Delta}_i(t_i, a'_i, a_i, a_{-i}) = \max\{\Delta_i(t_i, a'_i, a_i, a_{-i}), 0\} \). Now, suppose \( \sigma \) is not a minimax regret equilibrium. Let \( t_i \) and \( a_i \in A_i(\sigma_i(t_i)) \) be such that

\[
\max_{a_{-i} \in \Pi_{j \neq i}(\bigcup_{t_j \in T_j} A_j(\sigma_j(t_j)))} \tilde{\Delta}_i(t_i, a_i, a'_i, a_{-i}) > \max_{a_{-i} \in \Pi_{j \neq i}(\bigcup_{t_j \in T_j} A_j(\sigma_j(t_j)))} \tilde{\Delta}_i(t_i, a'_i, a_i, a_{-i}).
\]

Since \( \bigcup_{t_j \in T_j} A_j(\sigma_j(t_j)) = A_j, \forall j \neq i \), the above is equivalent to

\[
\max_{a_{-i} \in A_{-i}} \tilde{\Delta}_i(t_i, a_i, a'_i, a_{-i}) > \max_{a_{-i} \in A_{-i}} \tilde{\Delta}_i(t_i, a'_i, a_i, a_{-i}).
\]

Let \( \hat{a}_{-i} \) be such that \( \tilde{\Delta}_i(t_i, a_i, a'_i, \hat{a}_{-i}) = \max_{a_{-i} \in A_{-i}} \tilde{\Delta}_i(t_i, a_i, a'_i, a_{-i}) \). Consider \( \sigma^1 = (\sigma^1, \sigma_{-i}) \), where \( \sigma^1_i \in \Sigma_i(\sigma) \) is such that \( \sigma^1_i(t_i, a'_i) > \sigma_i(t_i, a_i) \) and \( \sigma^1_i(t_i, a_i) < \sigma_i(t_i, a_i) \). In state \( \sigma \), consider the ex-post event \( (t_i, t_{-i}) \) such that player \( i \) plays \( a_i \) and other players play \( \hat{a}_{-i} \in A_{-i} = \prod_{j \neq i} \left( \bigcup_{t_j \in T_j} A_j(\sigma_j(t_j)) \right) \). Then playing \( a'_i \) is a better reply to \( \hat{a}_{-i} \) than \( a_i \) and player \( i \)'s ex-post regret is \( \Delta_i(t_i, a_i, a'_i, \hat{a}_{-i}) > 0 \). Hence, \( \sigma^1 \in R_G(\sigma) \) and \( r(\sigma, \sigma^1) \leq \frac{1}{\Delta_i(t_i, a_i, a'_i, \hat{a}_{-i})} < \frac{1}{\gamma} \) since only a single player of type \( t_i \) in population \( \mathcal{P}_i \) switches her action from \( a_i \) to \( a'_i \) and there exists an ex-post event in state \( \sigma \) in which this player has positive ex-post regret of \( \Delta_i(t_i, a_i, a'_i, \hat{a}_{-i}) \) from playing \( a_i \) instead of \( a'_i \). Thus, in the graph \( W_G \), there exists an outgoing edge from \( \sigma \) to \( \sigma^1 \) with a weight of less than \( \frac{1}{\gamma} \). Now, consider any weighted cycle containing this outgoing edge. Since \( \sigma_i^1(t_i, a'_i) > \sigma_i(t_i, a'_i) \), it must be that there exist two states \( \sigma^l, \sigma^{l+1} \) in this weighted cycle such that \( \sigma^{l+1} \in R_G(\sigma^l) \) and \( \sigma^{l+1}(t_i, a'_i) < \sigma^l(t_i, a'_i) \). That is, there exists a player of type \( t_i \) in population \( \mathcal{P}_i \) who switches her action from \( a'_i \) to \( a_i \). However, \( \Delta_i(t_i, a_i, a'_i, \hat{a}_{-i}) = \tilde{\Delta}_i(t_i, a_i, a'_i, \hat{a}_{-i}) > \max_{a_{-i} \in A_{-i}} \tilde{\Delta}_i(t_i, a'_i, a_i, a_{-i}) \)

37
implies that $r(\sigma', \sigma^{l+1}) > \frac{1}{\Delta(t, a_i, a'_i, a_{i-1})}$. Thus, the state $\sigma$ passes the weighted cycle test, and hence it is not stochastically stable. \hfill \Box

**Appendix**

In this appendix, we briefly discuss stochastic stability and related concepts. For more details the reader is referred to the Appendix in Young (1993). A *finite Markov process* is a discrete time stochastic process over a finite number of states $S = \{s_1, \ldots, s_n\}$ defined by a matrix of transition probabilities

$$
\mathcal{M} = \begin{pmatrix}
    p_{11} & p_{12} & \cdots & p_{1n} \\
p_{21} & p_{22} & \cdots & p_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
p_{n1} & p_{n2} & \cdots & p_{nn}
\end{pmatrix}
$$

Here $p_{kl}$ is the probability of transitioning from state $s_k$ to state $s_l$ in one period. Note that we are considering only time homogeneous Markov processes since the transition probabilities do not depend on time.

An *invariant distribution* is a probability distribution $\mu = (\mu(s_1), \ldots, \mu(s_n))$ over $S$ such that

$$
\mu \mathcal{M} = \mu.
$$

It is well known that every finite Markov process has an invariant distribution.

A *recurrent class* is a set of states such that if the process reaches one of them, it will never leave the set, and such that it does not admit a proper subset of states with the same property. An *absorbing state* is a singleton recurrent class. A finite Markov process is *irreducible* if it has a unique recurrent class that is equal to the set of states. Let $T(s)$ be the set of integers $t \geq 1$ such that there is a positive probability of moving from state $s$ to state $s$ in exactly $t$ periods. A finite Markov process is *aperiodic* if 1 is the greatest common divisor of integers in $T(s)$ for all states $s$.

Let $\nu_t(s|s')$ denote the probability that the process is in state $s$ in period $t$ conditional on it being in state $s'$ in period 0. Let $\mu_t(s|s')$ be the relative frequency with which the process is in state $s$ in the first $t$ periods conditional on it being in state $s'$ in period 0. If a finite Markov process is both irreducible and aperiodic, then it

38
has a unique invariant distribution $\mu$ such that for all states $s$,

$$\mu(s) = \lim_{t \to \infty} \nu_t(s|s') = \lim_{t \to \infty} \mu_t(s|s'), \forall s'.$$

Thus, for an irreducible and aperiodic Markov process, irrespective of its initial state, the unique invariant distribution of the process approximates both the relative frequency of visiting each state over the long run and the probability of being in each state at any time period $t$ for sufficiently large $t$.

A finite Markov process $M^\epsilon$ is a regular perturbation of $M$ if (a) $M^\epsilon$ is irreducible and aperiodic for all $\epsilon \in (0, \epsilon^*]$, and for all $k, l = 1, \ldots, n$, we have (b) $\lim_{\epsilon \to 0} p_{kl}^\epsilon = p_{kl}$ (here $p_{kl}^\epsilon$ is the probability of transitioning from state $s_k$ to $s_l$ in one period under $M^\epsilon$) and (c) $p_{kl}^\epsilon > 0$ for some $\epsilon > 0$ implies that there exists an $r(s_k, s_l) \geq 0$ such that $0 < \lim_{\epsilon \to 0} \epsilon^{-r(s_k, s_l)}p_{kl}^\epsilon < \infty$. The number $r(s_k, s_l)$ is called the resistance of transitioning from state $s_k$ to $s_l$. Notice that $r(s_k, s_l)$ is unique for any two states $s_k$ and $s_l$, and $r(s_k, s_l) = 0$ if and only if $p_{kl} > 0$. For simplicity, we assume that $r(s_k, s_l) = \infty$ if $p_{kl}^\epsilon = 0, \forall \epsilon \in (0, \epsilon^*]$. Since $M^\epsilon$ is irreducible and aperiodic, it has a unique invariant distribution $\mu^\epsilon$.

A state $s$ is stochastically stable if $\lim_{\epsilon \to 0} \mu^\epsilon(s) > 0$.

Let $S_1, \ldots, S_K$ be the recurrent classes of $M$. A path from any recurrent class $S_k$ to $S_l$ is a finite sequence of states $s^1, \ldots, s^Z$ such that $s^1 \in S_k$ and $s^Z \in S_l$. The resistance of this path is defined as $r(s^1, s^2) + r(s^2, s^3) + \ldots + r(s^{Z-1}, s^Z)$. The resistance $r(S_k, S_L)$ is the minimum resistance over all paths from $S_k$ to $S_L$.

For any recurrent class $S_k$, an $S_k$-rooted tree is a weighted directed graph with $K$ vertices, equal to the set of recurrent classes $S_1, \ldots, S_K$, and $K - 1$ directed edges, one outgoing from each $S_l \neq S_k$, such that there is a unique directed path from each $S_l \neq S_k$ to $S_k$ and the weight of the directed edge from any $S_l$ to $S_l$ is equal to $r(S_l, S_l)$. The resistance of any $S_k$-rooted tree is equal to the sum of the weights on all its $K - 1$ edges. The stochastic potential of any recurrent class $S_k$ is equal to the minimum resistance over all $S_k$-rooted trees. It turns out that stochastically stable states are precisely those states that are contained in recurrent classes of $M$ that have the minimum stochastic potential.

**Theorem 7.1 (Young, 1993).** Let $M^\epsilon$ be a regular perturbation of $M$ and $\mu^\epsilon$ be its unique invariant distribution. As $\epsilon \to 0$, $\mu^\epsilon$ converges to an invariant distribution $\mu$ of $M$. Moreover, a state $s$ is stochastically stable if and only if $s$ is contained in a recurrent class of $M$ that has the minimum stochastic potential.
References


