Repeated Games with Forgetful Players*

Selçuk Özyurt †
New York University

December 7, 2012

Abstract

We present a model to investigate the behavior of forgetful players in infinitely repeated games. We assume that each player may forget the entire history of the play with a fixed probability. Our modeling specifications make a clear distinction between absentminded and forgetful players. We consider two extreme cases regarding the correlation of forgetfulness of the players. In the first case, forgetfulness is simultaneous: If a player forgets, so do the rest. For this part, we are able to prove two Folk theorems. In the other extreme, we consider the case where forgetfulness is independent between players, so players' state of memory is no longer common knowledge. We focus on *Conditionally Belief-Free* strategies to recapture the recursive structure in the sense of Abreu, Pearce and Stacchetti (1986, 1990). By utilizing a method analogous to Ely, Horner and Olszewski (2005), we represent characterization results for the payoff set of conditionally belief-free strategies.

*I owe special thanks to my advisor David Pearce for his valuable comments and suggestions at all stages of this project. I also thank the participants at Micro Student Workshop (2006, New York University), the 17^{th} International Conference on Game Theory, (2006, Stony Brook), and 2006 South and South East Asia Econometric Society Meeting (Chennai, India).

†Department of Economics, New York University, 19 W. 4^{th} Street, 6^{th} FL. New York, NY 10012 (e-mail: selcuk@nyu.edu)

1 Introduction

In this paper, we consider infinitely repeated games with forgetful players. The structure of each period in the repeated game is as follows: Players simultaneously choose their actions and then immediately these choices become publicly observable. Before moving to the next period, nature makes its move and sends one of two signals to each player: *Forget* with probability p, or *Recall*. Whenever a player receives the signal *Forget*, we say this player forgets the entire history of the play i.e., he forgets whatever he had chosen and observed previously. However, he can perfectly recall what period he is in. If, a player receives the *Recall* signal, he continues to perfectly recall all the actions and signals he has observed since the last period that he received a *Forget* signal. We assume that the probability of receiving the signal *Forget* is the same throughout the game and across players.

Our modeling assumptions make a clear distinction between forgetful players and absentminded players since even if players forget, they always remember the calendar time. As it is first introduced by Piccione and Rubinstein (1997), an absentminded player may visit the same information set more than once throughout the game. Remembering the calendar time makes such occurrences impossible in th game we consider. The reason for making this key assumption is to keep our analysis aside from the complications that might arise with absentminded players.[1]

For dynamic games, the perfect recall assumption simplifies the analysis of the game with a great deal. However, this assumption may be very strong and restrictive in some environments. The cognitive boundaries of the players may prevent them from perfectly recalling the history of the game. The decision-makers of a firm can change through time yet the firms continue to play the same game. In such instances, the payoff-relevant past play of the game may not be transferred through all generations perfectly, and the possibility of such an imperfect information flow between the decision-makers may affect their overall behavior.

Though forgetfulness can be modeled in various forms, we model it in a simple way. Yet, the complications following such a simple model are impressive. Strategies that have been used to prove various folk theorems in the repeated game literature are no longer equilibrium strategies in our context, because these strategies depend heavily

[1] More detailed discussions about absentminded players can be found at Piccione and Rubinstein (1997) and its references.

upon the assumption that each player can perfectly recall private or public history.

We consider two extreme cases in terms of the correlation of signals that players receive at each period. The first and relatively much simpler case assumes that the nature sends the same signal (perfectly correlated signals or simultaneous forgetfulness) to all players and this signal can be publicly observable. This assumption ensures that at any period of the repeated game, the players' state of memory will be the same and is common knowledge. In the second extreme case, we assume that the nature may send different signals (uncorrelated signals or independent forgetfulness) to each player and players privately observe these signals.

The latter case yields various complications that we do not encounter in the former case: At any period, the players' state of memory need not be same and this is private information. We also assume that players do not have an outside tool to communicate during the play of the repeated game. Therefore, each player must condition his strategies on his own private histories.

This paper is not the first attempt to analyze the behavior of players who have memory imperfections in a repeated game environment. Cole and Kocherlakato (2005) and Lehrer (1998) consider players who are capable of remembering only a fixed finite number of periods of past observations. In other words, in their set up, all the players are restricted to use strategies that depend only on fixed and finite length histories. These studies are analogous to our first case (simultaneous forgetfulness). But the main distinction with their set up and our first case is that our players' memory imperfection is stochastic i.e., there is always positive probability that all the players will remember the entire history of the game, this probability converging to zero as the game unfolds. Therefore, players are not restricted to strategies that depend on exogenously fixed finite length histories.

The second case (independent forgetfulness) is somehow more interesting but also more challenging. A Player may forget the history of the game and he never knows whether his opponents have forgotten and if not, what they can recall. So in this case each player's state of memory is not common knowledge. The main difficulty arising in this case is that players must condition their strategies on their private histories. Therefore, at any period the optimal continuation strategy for player i must be a best reply to his beliefs about his opponents' possible private histories. That is, the optimal continuation strategy is a correlated equilibrium where the correlation device is each player's beliefs about his opponents' private histories.

On the one hand, as the game unfolds this correlation device becomes more complicated. So, keeping track of these beliefs along each history and across players makes the verification of optimality of the continuation strategies an intractable problem. On the other hand, since the continuation strategies form correlated equilibrium, they need not be the equilibrium strategy of the original repeated game. In other words, we loose the recursive structure of the repeated game in the sense of Abreu, Pearce and Stacchetti (1986, 1990) (APS).

The independent forgetfulness case shares some common features with the repeated games with imperfect private monitoring. In both environments, players' strategies map private histories to action sets. Piccione (2002), Ely (2002), and Ely, Horner and Olszewski (2005) restrict players to play belief-free strategies; after any private history, a player's continuation strategy is a best reply to his opponents' continuation strategies regardless of the private histories each may have. These strategies are the "safest" strategies that players can use; even if players have the option to learn the true private histories of the game, they do not have incentive to change their own strategies. Ely et al. study this family of strategies in detail and show that the recursive structure of the repeated game can be recaptured if players are restricted to these strategies. However, this family is so restricted that it is not possible to support all feasible and individually rational payoffs as an equilibrium.[2]

In this paper, we show that by restricting players to use a class of strategies analogous to belief-free strategies (which we call conditionally belief-free strategies), it is possible to recapture the recursive structure in the sense of APS. A conditionally belief free strategy basically imposes the condition that for any player i and after any private history of this player, his continuation strategy is a best reply to his opponents' continuation strategies, as long as his opponents' private histories do not contradict the memory of player i's history. Note that in the game we consider, each player has some partial memory about the past play with a positive probability and the belief-free strategies do not allow players to use their partial memory. For this reason, the set of belief-free strategies is a subset of the set of conditionally belief-free strategies in this set-up.

So we relax the belief-freeness assumption by restricting players to use conditionally belief-free strategies so that they can utilize any information that they can recall. For example, suppose that player i recalls perfectly the entire history of the game at some

[2] As Ely, Horner and Olszewski (2005) show, the prisoners' dilemma is an exception i.e. every feasible and individually rational payoff of the prisoners' dilemma game can be supported as a belief-free equilibrium strategy.

period t i.e., he can perfectly recall all the action profiles played at each previous period. While the player i chooses his best reply continuation strategy at stage $t+1$, he does not know whether his opponent can perfectly recall this history or not, but if he recalls the entire history, player i knows that his opponent's private history will not be different than what he recalls, so player i does not have to select a continuation strategy which is best reply to his opponent's continuation strategy that follows a history "different" than what player i recalls (different in terms of action profiles played at each period).

It is clear that compared with belief-free strategies, conditionally belief-free strategies constitute a larger class of strategies. Moreover, as the probability of forgetting approaches zero, each player can perfectly recall the entire history of the game with high probability, and hence there is a high probability that their continuation strategies will be a best reply to the "true" continuation strategy that their opponents actually use. In other words, this class of strategies may be rich enough to prove that every feasible and individually rational payoff vector can be sustained as an equilibrium at least when probability of receiving the forget signal tends to 0.

Section 2 presents the basics of the model. In section 3 we consider simultaneous forgetfulness. For this case, we give two folk theorems: Nash threat and minimax threat folk theorems. The strategy that we use to prove the first result is a modified version of the strategy used by Friedman (1971). The minimax threat folk theorem's proof uses the modified version of the strategy introduced by Fudenberg and Maskin (1986).

In section 4, we focus on two-player games. The generalization to N players is possible though the analysis becomes significantly harder. Following a parallel method used by Ely et al., we restrict our attention to conditionally belief-free strategies in which at each period t, and after any t-length histories, players are restricted to choose their mixed actions over a fixed set, the "regime prevailing at period t". We define self generation of payoff vectors, which is analogous to APS, and similar to Ely et al..[3]

However, to recapture the recursive structure in our set up we need a stronger concept (we call as strong self generation) which employs the idea of self generating collection of sets. With this stronger definition at hand, we are able to show that if a set is strongly self generated, then it is a subset of the set of payoffs of all conditionally belief free equilibrium strategies, and conversely the set of payoff vectors of all conditionally belief free strategies is also strongly self generating. Finally, we conclude in section 5.

[3] Ely, Horner and Olszewski (2005) call this condition strong self generation

2 The Model

The stage game, G, consists of n (≥ 2) players, and for each player $i = 1, 2, ...n$, a finite action set $\mathbf{A}_i$ and the payoff function $u_i : \mathbf{A} := \mathrm{X}_{i=1}^{n} \mathbf{A}_i \to \mathbb{R}$. The infinitely repeated game is the repetition of the stage game G for indefinite time. We denote by $G^{\infty}(\delta, p)$ the infinitely repeated game with stage game G, discount factor $\delta \in (0, 1)$ and probability of forgetting $p \in (0, 1)$. We assume that the discount factor δ and the probability of forgetting p is same for each player.

Each period of the repeated game consists of three stages. In the first stage, players simultaneously chose their actions (possibly mixed actions). In the second stage, each players' choice is revealed and players publicly observe the revealed actions. In the final stage, nature either sends the signal *Forget*, with probability p, or the signal *Recall*, with probability $1 - p$, to each player. So, we model forgetfulness as a stochastic and exogenous action taken by nature. If the nature sends the signal *Forget* to player i at stage t, we say player i forgets the first t periods of the repeated game. So, every t length history of player i with which player i receives the signal *Forget* in period t will be collected into the same information set in period $t + 1$. If player i receives the signal *Recall*, he recalls all the information that he observed since the last time that he received the signal *Forget*. So, receiving the *Recall* signal in period t does not reveal any information about previous periods that are forgotten previously.

An n-tuple vector $a \in \mathbf{A}$ denotes a profile of actions, and $a_{-i} = (a_1, ..., a_{i-1}, a_{i+1}, ..., a_n)$ denotes the action profile of all players but i. A mixed action α_i for each player i is a randomization over $\mathbf{A}_i$ i.e., $\alpha_i \in \Delta\mathbf{A}_i$. Similarly $\alpha_{-i} = (\alpha_1, ..., \alpha_{i-1}, \alpha_{i+1}, ..., \alpha_n)$ denotes the mixture profile of all players but i. Player i's payoff if he takes action a_i given that others play according to α_{-i} is $u_i(a_i, \alpha_{-i})$.

For each player i, let $m^i = (m_1^i, ...m_n^i) \in \Delta\mathbf{A} := \Delta\mathbf{A}_1 \times ... \times \Delta\mathbf{A}_n$ so that

$$(m_1^i, ..., m_{i-1}^i, m_{i+1}^i, ..., m_n^i) = \arg\min_{\alpha_{-i}} \max_{a_i} u_i(a_i, \alpha_{-i})$$

and $v_i^* \equiv \max_{a_i} u_i(a_i, m_{-i}^i) = u_i(m^i)$.

The strategy profile $(m_1^i, ..., m_{i-1}^i, m_{i+1}^i, ..., m_n^i)$ is a minimax strategy (which may not be unique) against player i. v_i^* is the smallest payoff that the other players can keep player i below.[4] We call v_i^* player i's minimax value and refer to $(v_1^*, ..., v_n^*)$ as the

[4]Note that if the number of players is more than 2, other players may keep the player j's payoff even

minimax point of the game G. It is clear that in any equilibrium of the stage game G, whether or not it is repeated, player i's expected average payoff must be no less than v_i^*.

Henceforth, we normalize the payoffs of the stage game G in such a way that the minimax point is equal to the zero vector, i.e. $(v_1^*, ..., v_n^*) = (0, ..., 0)$. We denote $V = \textit{convex hull}\{v | \exists a \in A \text{ with } u(a) = v\}$ and $V^* = \{v \in V | v_i > 0 \text{ for all } i\}$.

The set V consists of feasible payoffs, and V^* consists of feasible and individually rational payoffs. In the repeated game $G^\infty(\delta, p)$ we assume that players maximize the discounted sum of single period payoffs. Therefore, if $\{a(t)\}_{t=1}^\infty$ is the sequence of vectors of actions played throughout the game, then player i's payoff is

$$\sum_{t=1}^{\infty} \delta^{t-1} u_i(a(t))$$

and the average payoff is

$$(1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(a(t))$$

We denote the set of signals by $\Sigma = \{\mathbf{F}, \mathbf{R}\}$ where $\mathbf{F}$ stands for *Forget* and $\mathbf{R}$ stands for *Recall*. A generic element of the set Σ is denoted by y, and y_i is the signal sent to player i. Let $\pi : \Sigma \to [0, 1]$ be the probability distribution over the set of signals where $\pi(y)$ is the probability of receiving the signal $y \in \Sigma$. Therefore, if $y = \mathbf{F}$, for instance, we have $\pi(y) = p$.

In the following two sections we will explore two extreme cases regarding the correlation between the signals that each player receives. Suppose that player i receives the signal $y_i \in \Sigma$ and the player i' receives the signal $y_{i'} \in \Sigma$. Section 3 considers the case where the correlation between y_i and $y_{i'}$ for any player i, i' is one, i.e. the signals are perfectly correlated. Section 4 explores the case where the correlation between the signals y_i and $y_{i'}$ for any player i, i', is zero: Players' signals are uncorrelated.

lower than v_j^* by using a correlated strategies against player j, where the outcome of the correlation device is not observed by j. Parallel to the rest of the literature on repeated games, however, we rule out such correlated strategies.

3 Perfectly Correlated Signals: Simultaneous Forgetfulness

In this section we assume that in each period the nature sends same signal to all players and that this signal is publicly observable. Therefore, players' state of memories is always same and this is common knowledge. In this environment, we are able to show that if the players are patient enough and if the probability of receiving the signal **F** is small enough, then every feasible and individually rational payoff vector is a payoff vector of some Perfect Bayesian Equilibrium (PBE) strategy profile.[5]

Proposition 3.1. *Let α^* be an equilibrium of the stage game with payoffs $y = (y_1, ...y_n)$. Then, for any $v = (v_1, ..., v_n) \in V$ with $v_i > y_i$ for all player i, there exists $x \in (0, 1)$ such that for all $\delta, p \in (0, 1)$ with $x < \delta(1 - p)$, there is a Perfect Bayesian Equilibrium of $G^\infty(p, \delta)$ with payoff v.*

Proof. See Appendix. □

In the proof of Proposition 3.1 we use a modified version of the strategy presented by Friedman (1971): For any $v = (v_1, ..., v_n) \in V$, players start playing the action profile that corresponds to the payoff vector v.[6] Call this strategy profile say $\hat{a} = (\hat{a}_1, ..., \hat{a}_n)$. Each player i continues to play $\hat{a}_i$ as long as the realized actions were $\hat{a}$ in all previous periods that all players can remember. If one player deviates from $\hat{a}$, and if players can recall this deviation, each player i starts playing α_i^*-the stage game Nash Equilibrium strategy- for the rest of the game. When players receive the signal **F**, they start playing according to $\hat{a}$.

Proposition 3.1 is silent for the individually rational and feasible payoffs that are strictly less than the static Nash Equilibrium payoff of the stage game G. In order to cover those payoffs, we need more complex strategies. The next result generalizes Proposition 3.1.

[5] We use Perfect Bayesian Equilibrium as the equilibrium concept in this paper. Sequential equilibrium, as it is introduced by Kreps and Wilson (1982), is defined for finite dynamic games with perfect recall. It is still an open question which equilibrium notion would be most suitable for games with imperfect recall.

[6] The payoff vector v possibly does not correspond to a pure strategy action profile $\hat{a}$. In that case $\hat{a}$ is a public randomization yielding payoffs with expected value v.

Proposition 3.2. *Assume that the dimension of the set V^* of feasible and individually rational payoffs equal the number of players. Then, for any $v = (v_1, ..., v_n) \in V^*$, there exists $x \in (0, 1)$ such that for all $\delta, p \in (0, 1)$ with $x < \delta(1-p)$ there is a Perfect Bayesian Equilibrium of $G^\infty(p, \delta)$ with payoff v.*

Proof. See appendix. □

In the proof of Proposition 3.2 we use a modified version of Fudenberg and Maskin (1986). We fix some $v \in V^*$. Each player i starts playing $\widehat{a}_i$ where $u(\widehat{a}) = v$.[7] Players continue playing according to $\hat{a}$ as long as no player deviates. If a player deviates, he is minimaxed by the other players for a finite period that is long enough to remove any gain from deviating from the initial strategy. Once the punishment phase is over, players move to the reward phase for the rest of the game, i.e. players start playing according to $\hat{\hat{a}}$ that "rewards" only the minimaxing players in the form of an additional "ϵ " in their average payoff. The reason for that is to give the required incentive for the players to minimax the deviating player in the punishment phase. The requirement that the dimension of the set V^* is equal to the number of players ensures the existence of such a reward in the payoff set. When players receive the signal **F**, whether they are in the punishment phase or in the reward phase, they revert to playing their initial actions according to $\hat{a}$. It is important to note that Proposition 3.1 and 3.2 are proved under the assumption that there exists a public randomization device, and that mixed strategies are observable. The proofs can be replicated even if the players do not observe the mixed actions.

4 Uncorrelated Signals: Independent Forgetfulness

In this section, we assume that the nature sends (possibly different) signals to each player and players privately observe their signals. Therefore players' states of memories do not have to be same, and they are not common knowledge. That is, if player i receives the signal **F**, he does not know whether his opponent received the same signal or not. All he knows is that his opponent might have received the signal **F** with probability p or the signal **R** with probability $1 - p$.

The independent forgetfulness case is rather difficult compared with the perfectly correlated case. The difficulty arises because players need not agree on the "true history"

[7] $\hat{a}$ may be a public randomization yielding payoffs with expected value v.

of the game: If player i forgets, he does not know which history is the true history of the repeated game. Moreover, he does not know whether his opponent has also forgotten, or if recalls, which periods he recalls. Therefore, in this case, players' strategies must depend on their private histories.

In addition, in any period and after any private history, player i's best reply must be based on an expected payoff calculation relative to i's beliefs about his opponents' possible histories. This implies that at each period the optimal continuation strategy must be a correlated equilibrium in which the correlation device is the players' beliefs about their opponents' possible histories. Therefore, after any private history the continuation strategy need not be the equilibrium strategy of the original game and so we loose the recursive structure in the sense of APS.

On the other hand, as the game unfolds, the correlation device becomes much more complicated, and the need of keep track of beliefs of players over time and across histories makes the verification of optimality of the continuation strategy an intractable exercise.

In this section we show that if we restrict our attention to a specific family of strategies, it is possible to recapture the recursive structure. The family of strategies we consider is called *Conditionally Belief-Free*.

4.1 Preliminaries

In this section, we focus on two-player games. $\mathbf{A} = \mathbf{A}_1 \times \mathbf{A}_2$ is the set of action profiles. A t-length (private) history for player i is an element of $H_i^t := (\mathbf{A} \times \Sigma)^t$. Given that $a \in \mathbf{A}$ and $y_i \in \Sigma$ are generic elements of $\mathbf{A}$ and Σ respectively, let $h_i^t = \{(a^\tau, y_i^\tau)\}_{\tau=1}^{\tau=t}$ be a generic element of H_i^t. If a player's initial history is the null history, we denote it $\emptyset$.

A t-length history for player i is called history with memory k (where $1 \leq k \leq t$) if the signal $\mathbf{F}$ is not observed for last k periods. That is, the history $h_i^t = \{(a^\tau, y_i^\tau)\}_{\tau=1}^{\tau=t}$ has memory k if $y_i^\tau = \mathbf{R}$ for every τ with $t-k+1 \leq \tau \leq t$. In case the t-length history h_i^t is specified to have a particular memory, say k, we write it as $h_i^t(k)$.

For any t-length, k-memory history, $h_i^t(k)$, we call the sequence of the last k period actions, $m^{h_i^t(k)} = (\{a^\tau\}_{\tau=t-k+1}^{\tau=t}, t)$, the memory of $h_i^t(k)$. Note that the order of a^τ's in a memory is important although the notation does not explicitly emphasize this point. If $k = 0$, we take $m^{h_i^t(k)} = (\varnothing, t)$. Two histories with same memory are treated differently if the length of these histories is different, so we define the memory as a sequence of actions and the total length of the history.

Let $H_i^t(k)$ be the set of i's private t-length, k-memory histories. For every t-length, k-memory history, $h_i^t(k)$, define $\overline{h_i^t(k)}$ to be the partition of the set $H_i^t(k)$, where $\overline{h_i^t(k)}$ is the set of all t-length, k-memory histories that agrees with the memory of $h_i^t(k)$. More formally, for any $h_i^t(k) = \{(a^\tau, y_i^\tau)\}_{\tau=1}^{\tau=t}$, the history $\hat{h}_i^t(k) = \{(\hat{a}^\tau, \hat{y}_i^\tau)\}_{\tau=1}^{\tau=t}$ belongs to $\overline{h_i^t(k)}$ if and only if $\hat{a}^\tau = a^\tau$ for each τ satisfying $t - k + 1 \leq \tau \leq t$.

For every t-length, k-memory history, $h_i^t(k)$, define the equivalence class $\widetilde{h_i^t(k)}$ to be the set of all t-length histories that agrees with the memory of $h_i^t(k)$. That is, if $h_i^t(k) = \{(a^\tau, y_i^\tau)\}_{\tau=1}^{\tau=t}$, a t-length history (with any memory) $\hat{h}_i^t = \{(\hat{a}^\tau, \hat{y}_i^\tau)\}_{\tau=1}^{\tau=t} \in \widetilde{h_i^t(k)}$ if and only if $a^\tau = \hat{a}^\tau$ for every τ which satisfies $t \geq \tau \geq t-k+1$. Note that any t-length history with 0-memory has $\widetilde{h_i^t(0)} = H_i^t$. Also note that histories in $\widetilde{h_i^t(k)}$ depend only on the memory of the history $h_i^t(k)$, i.e., not the past observation of signals.

Let $H_i = \cup_{t,k} H_i^t(k)$ be the set of private histories for player i. Define H as the set of history profiles i.e., $H = \{h^t = \{(a^\tau, y_i^\tau), (\hat{a}^\tau, \hat{y}_{-i}^\tau)\}_{\tau=1}^{\tau=t} \in H_i \times H_{-i} | t \in \mathbb{Z}_+$ such that $a^\tau = \hat{a}^\tau$ for every $t \geq \tau \geq t - \max\{k, \hat{k}\} + 1$ where $k, \hat{k}$ are the memories of h_i^t and h_{-i}^t respectively$\}$. Note that $H \neq H_i \times H_{-i}$. A pair of t-length histories is denoted by h^t.

For any history profile h, let $s|_h$ denote the continuation strategy profile derived from s following history h. Given a strategy profile s, for each t and $h_{-i}^t \in H_{-i}^t$, let $B_i(s|_{h_{-i}^t})$ denote the set of continuation strategies for player i that are best replies to $s_{-i}|_{h_{-i}^t}$.

We give a generic definition of the (behavior) strategy in our context.

Definition 4.1. *A repeated game (behavior) strategy for player i is a mapping s_i : $H_i \longrightarrow \Delta A_i$ with the property that for any t, k, and any different $h, h' \in \overline{h_i^t(k)}$, $s_i(h) = s_i(h')$.*

According to Ely et al., a strategy profile s is belief-free if for every $h_i^t \in H_i$, $s_i|_{h_i^t} \in B_i(s|_{h_{-i}^t})$ for all $h_{-i}^t \in H_{-i}^t$, and for $i = 1, 2$. It is easy to see that belief-free strategies are PBE. Moreover, they are the "safest" strategies that players can use: A player's continuation strategies must be best replies to opponents' continuation strategies regardless of the private history he might have. So, if both players use belief-free strategies, during the repeated game, they will never change their strategies even if they have a chance to learn the true history.

Ely et al. show that it is possible to recapture the recursive structure if we concentrate on this specific family of strategies. However, this family is not large enough to attain folk theorems for many games. Obviously, we can use this family of strategies in our set up as well; however, we show that the recursive structure in our setup can be recaptured by a more general class of strategies.

Definition 4.2. *A strategy profile s is conditionally belief-free if for every $h_i^t(k) \in H_i$, $s_i|_{h_i^t(k)} \in B_i(s|_{h_{-i}^t})$ for all $h_{-i}^t \in \widetilde{h_i^t(k)}$ and for $i = 1, 2$.*

A strategy profile is *conditionally belief-free* if and only if for any t-length, k-memory history, $h_i^t(k)$, player i's best reply strategy is independent of player $-i$'s private histories, conditional on that $-i$'s private histories do not contradict with the memory of the history $h_i^t(k)$. If, for instance, player i has a history of memory zero, then his best reply strategy is independent of other player's private histories. However, if i has a history of memory one, then his best reply strategy depends on the memory of that history: Since i remembers only the last stage of the repeated game, he does not have to choose his continuation strategy which is best reply to all possible t-length histories.

By using conditionally belief-free strategies, we relax the strong assumption made by belief-free strategies. In our case, forgetful players may have partial memories, yet belief-free strategies do not use this partial information. Therefore, we allow players to use any piece of information that can be recalled.

It is important to note that conditionally belief-free strategies are not as strong as they seem. For instance, subgame perfect equilibrium strategies in repeated games with perfect monitoring are conditionally belief-free when there is perfect recall. Similarly, sequential equilibrium or perfect public equilibrium strategies in games with imperfect but public monitoring are also conditionally belief-free as long as there is perfect recall. This is because players' private histories coincide in all these cases.

It is clear that conditionally belief-free strategies are PBE. Perfect Bayesian Equilibrium strategies involving conditionally belief-free strategies will be called conditionally belief-free equilibrium.

Continuation Values

For any conditionally belief-free equilibrium strategy profile, the continuation strategy after some history profile $h^t \in H^t$ is also conditionally belief-free. Although each player might disagree on the history of the play in h^t, each player's continuation strategy after the history h^t is same as the continuation strategy which follows the true history, by definition of conditionally belief-freeness.

Moreover, for any conditionally belief-free strategy and player there is a set of potential continuation values. These values are the payoffs from the continuation strategy profiles that could arise after some finite history of the play. Given player $-i$'s history h_{-i}^t, the continuation value of player i, $u_i(s|\, h_i^t)$ (for any h_i^t whose memory does not contradict with the memory of h_{-i}^t), can be treated as a function $w_i(a^t, y_{-i}^t)$. This function depends on the previous period's action profile a^t and player $-i$'s private signal y_{-i}^t. This is because given player $-i$'s private history, each continuation strategy played by player i after any history h_i^t (whose memory does not contradict with the memory of h_{-i}^t) is a best-reply and hence achieves the same value.[8][9]

Fix some set $W = W_1 \times W_2 \subseteq \mathbb{R}^2$. A continuation payoff function for player i is a function of the form $w_i : \mathbf{A} \times \mathbf{\Sigma}_{-i} \longrightarrow W_i$ with the following property: For any a, $a' \in \mathbf{A}$, $w_i(a, y_{-i}) = w_i(a', y_{-i})$ whenever $y_{-i} = \mathbf{F}$. This condition requires that given that the opponent receives the signal $\mathbf{F}$ in last period, it is irrelevant whether the player recalls the action profile a or a'. This condition is implied by the fact that players are restricted to use conditionally belief-free strategies.

Regimes

Let $\mathcal{P}(\mathbf{A}) = \mathcal{P}(\mathbf{A}_1) \times \mathcal{P}(\mathbf{A}_2)$ be the set of all non-empty subsets of $\mathbf{A}$, and let $\mathcal{A}_i \in \mathcal{P}(\mathbf{A}_i)$ for $i = 1, 2$. In case $\mathcal{A} = (\mathcal{A}_i, \mathcal{A}_{-i})$ is specified to have a particular period t, we write it as $\mathcal{A}^t$, and we call $\mathcal{A}^t$ the regime that prevails at date t. We restrict our attention to a specific family of conditionally belief-free equilibrium strategies:

[8] It is important to note that the continuation value function $w_i(.,.)$ depends on player $-i$'s history h_{-i}^{t-1} although the representation does not explicitly reveals this dependency.

[9] Fix the history h_{-i}^t and a conditionally belief-free strategy. Consider a history h_i^t whose memory does not contradict with the memory of h_{-i}^t. Suppose first that h_i^t's memory is "shorter" than the memory of h_{-i}^t. Then since the strategy is conditionally belief free, the continuation strategy following h_i^t is best reply to continuation strategy for player $-i$ following the history h_{-i}^t. However, if h_i^t's memory is "longer" than the memory of h_{-i}^t, then we can always find a history $\hat{h}_{-i}^t$ whose memory is exactly same as the memory of h_{-i}^t such that the history $\hat{h}_{-i}^t$ does not contradict with the memory of h_i^t. Since histories h_{-i}^t and $\hat{h}_{-i}^t$ have same memories, $-i$'s continuation strategy must be same after these two histories, and hence, player i's continuation strategy after history h_i^t is best reply to $-i$'s continuation strategy following the history h_{-i}^t.

Definition 4.3. *A strategy profile (s_i, s_{-i}) is called conditionally belief-free with full support over the regime sequence $\{\mathcal{A}^t\}$ if for any i, period t, and $h_i^t \in H_i^t$, we have $s_i(h_i^t) \in \Delta\mathcal{A}_i^t$, and $s_i(h_i^t)[a_i] > 0$ for all $a_i \in \mathcal{A}_i^t$.*

Obviously this is a restricted class of strategies, because there might be conditionally belief-free strategies that randomize over different sets when a player receives the signal **F** and when he receives the signal **R**. This restriction yields important simplifications for the rest of our analysis: The following Propositions and the Corollary will shape our definition of strong self-generation. Their proof is very similar to one given by Ely et al.

Proposition 4.1 (Exchangeability Property). *Let $\{\mathcal{A}^t\}$ be a sequence of regimes and let s, s' be two conditionally belief-free equilibria with full support over the regime sequence $\{\mathcal{A}^t\}$. The strategy profiles (s_1, s_2') and (s_1', s_2) are also conditionally belief-free equilibria with full support over the regime sequence $\{\mathcal{A}^t\}$.*

Proof. See Appendix. □

The following corollary is a straightforward application of Proposition 4.1.

Corollary 4.1. *Let $E(\{\mathcal{A}^t\})$ be the set of all payoffs arising from conditionally belief-free equilibria bounded by the regime sequence $\{\mathcal{A}^t\}$. Then $E(\{\mathcal{A}^t\}) = E_1 \times E_2$ for some subsets E_1, E_2 of $\mathbb{R}$.*

4.2 Strong Self Generation

In this section, we fix the value of the discount factor δ and probability of forgetting p where $\delta, p \in (0, 1)$.

Definition 4.4 (Self Generation). *Let $W = W_1 \times W_2 \subseteq \mathbb{R}^2$ and the regime $\mathcal{A} \in \mathcal{P}(\mathbf{A})$ be given. A payoff vector $v \in \mathbb{R}^2$ is generated by W using the regime $\mathcal{A}$ if for each player i there is a mixture $\alpha_{-i} \in \Delta\mathcal{A}_{-i}$ and a continuation payoff function $w_i : \mathbf{A} \times \Sigma_{-i} \to W_i$ such that*

$$v_i \geq (1-\delta)u_i(a_i, \alpha_{-i}) + \delta E(\, w_i |\, a_i, \alpha_{-i}, y_{-i})$$

for all $a_i \in \mathbf{A}_i$, with equality for each $a_i \in \mathcal{A}_i$, where

$$E(\, w_i |\, a_i, \alpha_{-i}, y_{-i}) = \sum_{a_{-i} \in \mathbf{A}_{-i}} \sum_{y_{-i} \in \Sigma} \alpha_{-i}(a_{-i})\pi(y_{-i})w_i(a_i, a_{-i}, y_{-i})$$

This definition is parallel to the self-generation of APS and same with the strong self-generation of Ely et al.

When there exists α_{-i} and w_i as in the definition, say α_{-i} enforces w_i and generates v_i, or (α_{-i}, w_i) generates v_i in brief. We fix the regime $\mathcal{A}$ throughout.

In this setup, self generation as given in previous definition is not enough to achieve a characterization result as in the case of APS. The reason for this is that, at each period, once player i receive the signal $\mathbf{F}$, his opponent's best reply set is independent of player i's private history. Therefore, player $-i$'s continuation value must achieve the same value independent of his private history. Therefore, at any period, the continuation value for player $-i$ following histories with zero memory for player i must be same. Self generation, however, does not impose any restriction for this event to be true. For this purpose, we consider the self generating sets of sets. So, an $\mathbf{F}$-generated set corresponds to the values of all continuation strategies of a conditionally belief-free strategy at some period t.

Definition 4.5 (F-generated sets). *Let $W = W_1 \times W_2 \subseteq \mathbb{R}^2$. For any player i, a set of real numbers $X \subseteq \mathbb{R}$ is $\mathbf{F}$-generated by W_i, if for each $v \in X$,*

1. *there exists a mixture $\alpha_{-i} \in \Delta\mathcal{A}$ and a continuation payoff function $w_i : \mathbf{A} \times \Sigma_{-i} \to W_i$ such that (α^v_{-i}, w^v_i) generates v, and*
2. *for any $v, v' \in X$, we have $w^v_i(., y_{-i}) = w^{v'}_i(., y_{-i})$ whenever $y_{-i} = \mathbf{F}$.*

Fix the set $W = W_1 \times W_2$ throughout. Let $\mathcal{P}(W_i)$ denote the set of all non-empty subsets of W_i and let $\overline{W}_i$ be a non-empty subset of $\mathcal{P}(W_i)$ in which each set $X \in \overline{W}_i$ is $\mathbf{F}$-generated by W_i. We call such collections of sets as $\mathbf{F}$-collection.

A set of real numbers $X \subseteq \mathbb{R}$ is called $\mathbf{F}$-generated by the $\mathbf{F}$-collection $\overline{W}_i$, if X is $\mathbf{F}$-generated by W_i and the set of continuation values $\{w^v_i(a, y_{-i}) \mid \exists a \in \mathbf{A}$, $y_{-i} \in \Sigma_{-i}$, $v \in X$, $\alpha^v_{-i} \in \Delta\mathcal{A}_{-i}$, and w^v_i s.t. (α^v_{-i}, w^v_i) generates v $\}$ is a subset of some set in $\overline{W}_i$. Define $\overline{B}(\overline{W}_i)$ as the set of all sets of real numbers that are $\mathbf{F}$-generated by the $\mathbf{F}$-collection $\overline{W}_i$. Therefore, an $\mathbf{F}$-collection, $\overline{W}_i$, is called *self-generating F-collection* if $\overline{W}_i \subseteq \overline{B}(\overline{W}_i)$. Finally, define $B^*(\overline{W}_i) = \{v \mid \exists X \in \overline{W}_i \text{ s.t. } v \in X\}$.

Definition 4.6. *Let $W = W_1 \times W_2 \subseteq \mathbb{R}^2$ be given. W is strongly self-generating if for each player i, $W_i \subseteq B^*(\overline{W}_i)$ for some self-generating $\mathbf{F}$-collection $\overline{W}_i$.*

The next proposition shows that if a set W is a strongly self-generating set using $\mathcal{A}$, then it is a subset of the payoff set of all conditionally belief-free equilibrium strategies with full support over the regime $\mathcal{A}$.

Proposition 4.2. *If W is strongly self-generating set using $\mathcal{A}$, then each element of W is payoff of a conditionally belief-free equilibrium strategy with full support over the constant regime $\mathcal{A}$.*

Proof. See Appendix. □

The next proposition corresponds to the factorization theorem of APS. It shows that the payoff set of all conditionally belief-free equilibrium strategies with full support over the fixed regime $\mathcal{A}$ is itself strongly self generating.

Proposition 4.3. *The payoff set of all conditionally belief-free equilibrium strategies with full support over the constant regime $\mathcal{A}$ is itself a strongly self-generating set.*

Proof. See Appendix. □

Previous propositions characterize conditionally belie-free equilibrium strategies using a fixed regime in every period. Obviously this is a restricted family of conditionally belief-free strategies, and there is room to construct more equilibrium payoffs: If we relax the single regime assumption and allow using possibly different regime at every period, we can attain more conditionally belief-free equilibrium payoffs. To pursue in this way, parallel to Ely et al., we use public randomization over regimes.

4.3 Public Randomization over Regimes

We, now, suppose that at each period, players publicly observe the outcome of a lottery over the set of regimes. Our interpretation is that when a certain regime $\mathcal{A}$ is realized in a certain period, players choose their optimal actions over the regime $\mathcal{A}$, more precisely the optimal actions must be full support over the regime $\mathcal{A}$. We assume that players never forget the history of public randomization outcome. Therefore, if a player receives the signal $\mathbf{F}$, then he forgets the history of the action-signal profiles, but perfectly recalls the history of the public randomization. This assumption, we believe, is without loss of generality, yet allowing players the possibility of forgetting the public randomization makes the analysis harder.

When players have access to a public randomization device, a strategy shall depend on private history as well as the history of realizations of the public randomization. We make the following modifications to our notations; Let $\mathbf{Z}$ be the set of public signals, and $\mathbf{z}$ be a generic element of $\mathbf{Z}$. We will assume direct public randomization, i.e., $\mathbf{Z} = \mathcal{P}(\mathbf{A})$. A t-length (private) history for player i is an element of $H_i^t := (\mathbf{A} \times \Sigma \times \mathbf{Z})^t$. So, $h_i^t = \{(a^\tau, y_i^\tau, \mathbf{z}^t)\}_{\tau=1}^{\tau=t}$ is a generic element of H_i^t.

For any t-length, k-memory history, $h_i^t(k)$, we call the sequence of the last k period actions and the sequence of public randomization realizations, $m^{h_i^t(k)} = (\{a^\tau\}_{\tau=t-k+1}^{\tau=t}, \{\mathbf{z}^t\}_{\tau=1}^{\tau=t}, t)$, memory of $h_i^t(k)$. If $k = 0$ we take $m^{h_i^t(k)} = (\{\mathbf{z}^t\}_{\tau=1}^{\tau=t}, t)$. For every t-length, k-memory history, $h_i^t(k)$, define the equivalence class $\overline{h_i^t(k)}$ be the set of all t -length, k-memory histories that agrees with the memory of $h_i^t(k)$. That is, if $h_i^t(k) = \{(a^\tau, y_i^\tau, \mathbf{z}^\tau)\}_{\tau=1}^{\tau=t}$, for any t-length, k-memory history $\hat{h}_i^t = \{(\hat{a}^\tau, \hat{y}_i^\tau, \mathbf{z}^\tau)\}_{\tau=1}^{\tau=t} \in \overline{h_i^t(k)}$ we have $a^\tau = \hat{a}^\tau$ whenever $t \geq \tau \geq t - k + 1$.

Similarly, for every t-length, k-memory history, $h_i^t(k)$, define the equivalence class $\widetilde{h_i^t(k)}$ be the set of all t-length histories that agrees with the memory of $h_i^t(k)$. That is, if $h_i^t(k) = \{(a^\tau, y_i^\tau, \mathbf{z}^\tau)\}_{\tau=1}^{\tau=t}$, for any t-length history (with any memory in which the history of realization of public randomization is same as $h_i^t(k)$) $\hat{h}_i^t = \{(\hat{a}^\tau, \hat{y}_i^\tau, \mathbf{z}^\tau)\}_{\tau=1}^{\tau=t}$ $\in \widetilde{h_i^t(k)}$ we have $a^\tau = \hat{a}^\tau$ for every τ which satisfies $t \geq \tau \geq t - k + 1$.

Let H is the set of history profiles defined as $H = \{h^t = \{(a^\tau, y_i^\tau, \mathbf{z}^\tau), (\hat{a}^\tau, \hat{y}_i^\tau, \mathbf{z}^\tau)\}_{\tau=1}^{\tau=t}$ $\in H_i \times H_{-i} \mid t \in \mathbb{Z}_+$ such that $a^\tau = \hat{a}^\tau$ for every $t \geq \tau \geq t - \max\{k, \hat{k}\} + 1$ where $k, \hat{k}$ are the memories of h_i^t and h_{-i}^t respectively$\}$.

Given that the public randomization is included into our notations this way, our definitions of (behavior) strategy and conditionally belief-free strategy remain the same. For any $\mathbf{z} \in \mathbf{Z}$, the regime in a given period depends on the current realization of $\mathbf{z}$, $\mathcal{A}^t(\mathbf{z}^t)$ where $\mathcal{A}^t(\mathbf{z}^t) = \mathcal{A}_i^t(\mathbf{z}^t) \times \mathcal{A}_{-i}^t(\mathbf{z}^t)$. Therefore, a strategy profile (s_i, s_{-i}) is called conditionally belief-free strategy with full support according to the *i.i.d.* public randomization $\mu \in \Delta\mathcal{P}(\mathbf{A})$ if for any i, period t, and $h_i^t \in H_i^t$, we have $s_i(h_i^t) \in \Delta\mathcal{A}_i^t(\mathbf{z}^t)$, and $s_i(h_i^t)[a_i] > 0$ for all $a_i \in \mathcal{A}_i^t(\mathbf{z}^t)$ where $\mathbf{z}^t$ is the realization of public randomization in h_i^t.

Strong self-generation is now defined with respect to a fixed public randomization over regimes. Let $\mu \in \Delta\mathcal{P}(\mathbf{A})$ be an *i.i.d.* probability distribution over the set of all non-empty regimes.

Definition 4.7. *Let $W = W_1 \times W_2 \subseteq \mathbb{R}^2$, and i.i.d. public randomization over regimes*

$\mu \in \Delta\mathcal{P}(\mathbf{A})$ *be given. A payoff vector* $v \in \mathbb{R}^2$ *is generated by* W *using the public randomization* μ *if for each player* i*, and regime* $\mathcal{A} \in \mathcal{P}(\mathbf{A})$*, there is a mixture* $\alpha_{-i}^{\mathcal{A}} \in \Delta\mathcal{A}_{-i}$ *and a continuation payoff function* $w_i^{\mathcal{A}} : \mathbf{A} \times \Sigma_{-i} \to W_i$ *such that, for each* $a_i : \mathcal{P}(\mathbf{A}) \to \mathbf{A}_i$ *we have*

$$v_i \geq \sum_{\mathcal{A} \in \mathcal{P}(\mathbf{A})} \mu(\mathcal{A}) \left[(1-\delta)u_i(a_i(\mathcal{A}), \alpha_{-i}^{\mathcal{A}}) + \delta E(w_i^{\mathcal{A}} \big| \, a_i, \alpha_{-i}^{\mathcal{A}}, y_{-i})\right] \tag{1}$$

for all $a_i \in \mathbf{A}_i$*, with equality for each* $a_i(\mathcal{A}) \in \mathcal{A}_i$*, where*

$$E(w_i^{\mathcal{A}} \big| \, a_i, \alpha_{-i}, y_{-i}) = \sum_{a_{-i} \in \mathbf{A}_{-i}} \sum_{y_{-i} \in \Sigma} \alpha_{-i}^{\mathcal{A}}(a_{-i})\pi(y_{-i})w_i^{\mathcal{A}}(a_i(\mathcal{A}), a_{-i}, y_{-i}) \tag{2}$$

When there exists $\alpha_{-i}^{\mathcal{A}}$ and $w_i^{\mathcal{A}}$ for each $\mathcal{A} \in \mathcal{P}(\mathbf{A})$ as in the definition, say the collection of mixtures $\{\alpha_{-i}^{\mathcal{A}}\}$ enforces the collection of $\{w_{-i}^{\mathcal{A}}\}$ respectively and generates v_i, or $(\{\alpha_{-i}^{\mathcal{A}}\}, \{w_i^{\mathcal{A}}\})$ generates v_i in brief.

Now we define self-generation of sets of real numbers. For that purpose, fix the set $W = W_1 \times W_2$ and the $i.i.d$ public randomization $\mu \in \Delta\mathcal{P}(\mathbf{A})$ throughout.

Definition 4.8. *For any player* i*, a set of real numbers* $X \subseteq \mathbb{R}$ *is* $\mathbf{F}$*-generated by* W_i *using* μ*, if for each* $v \in X$*, and each regime* $\mathcal{A} \in \mathcal{P}(\mathbf{A})$*,*

1. *there exists a mixture* $\alpha_{-i}^{\mathcal{A}} \in \Delta\mathcal{A}$ *and a continuation payoff function* $w_i^{\mathcal{A}} : \mathbf{A} \times \Sigma_{-i} \longrightarrow W_i$ *such that* $(\{\alpha_{-i}^{\mathcal{A},v}\}, \{w_i^{\mathcal{A},v}\})$ *generates* v*, and*
2. *for any* $v, v' \in X$ *and* $\mathcal{A} \in \mathcal{P}(\mathbf{A})$*, we have* $w_i^{\mathcal{A},v}(., y_{-i}) = w_i^{\mathcal{A},v'}(., y_{-i})$ *whenever* $y_{-i} = \mathbf{F}$.

Let $\overline{W}_i^{\mu}$ is a non-empty subset of $\mathcal{P}(W_i)$ in which each set $X \in \overline{W}_i^{\mu}$ is $\mathbf{F}$-generated by W_i using μ. We call such collections of sets as **F-**collection. A set of real numbers $X \subseteq \mathbb{R}$ is $\mathbf{F}$-generated by the $\mathbf{F}$-collection $\overline{W}_i^{\mu}$, if X is $\mathbf{F}$-generated by W_i and the set of continuation values $\{w_i^{\mathcal{A}^*,v}(a, y_{-i}) \mid \exists a \in \mathbf{A},\ y_{-i} \in \Sigma_{-i},\ \mathcal{A}^* \in \mathcal{P}(\mathbf{A})$ and $v \in X$, $\alpha_{-i}^{\mathcal{A},v} \in \Delta\mathcal{A}_{-i}$, $w_i^{\mathcal{A},v}$ for each $\mathcal{A}$ s.t. $(\{\alpha_{-i}^{\mathcal{A},v}\}, \{w_i^{\mathcal{A},v}\})$ generates v $\}$ is a subset of some set in $\overline{W}_i^{\mu}$.

Let $\overline{B}(\overline{W}_i^{\mu})$ is the set of all set of real numbers that are $\mathbf{F}$-generated by $\overline{W}_i^{\mu}$. An **F-**collection, $\overline{W}_i^{\mu}$, is called *self-generating$F$-collection* if $\overline{W}_i^{\mu} \subseteq \overline{B}(\overline{W}_i^{\mu})$. Define $B^*(\overline{W}_i^{\mu}) = \{v \mid \exists X \in \overline{W}_i^{\mu}$ s.t. $v \in X\}$.

Definition 4.9. *Let* $W = W_1 \times W_2 \subseteq \mathbb{R}^2$ *be given.* W *is strongly self-generating set using* μ *if for each player* i*,* $W_i \subseteq B^*(\overline{W}_i^{\mu})$*, for some self-generating* $\mathbf{F}-$*collection,* $\overline{W}_i^{\mu}$*.*

The following proposition is the analogue of Proposition 4.2.

Proposition 4.4. *If W is strongly self-generating set using the i.i.d. public randomization $\mu \in \Delta\mathcal{P}(\mathbf{A})$, then each element of W is the payoff of a conditionally belief-free strategy with full support according to μ.*

Proof. See Appendix. $\square$

The following proposition, however, is the analogue of Proposition 4.3.

Proposition 4.5. *The set of all payoffs of the conditionally belief-free equilibria with full support according to μ is itself a strongly self-generating set.*

Proof. See Appendix. $\square$

5 Conclusion

In this study, we analyze infinitely repeated games with forgetful players. The first case, where all players forget simultaneously, is rather simpler and so, we can provide folk theorems. The simplicity is due to the fact that the common knowledge assumption of the histories is preserved. In the second case, where players forget independently, players (private) histories are no longer common knowledge. We restrict players to use conditionally belief-free strategies that naturally depend only on players' own private histories. We are able to show that the recursive structure in the sense of APS can be recaptured in our case, however for this to be happen, we need a stronger notion of self generation; self generating sets of sets. We represent the analysis of this section for two players. The generalization to a finite N players is simple, however, the analysis are significantly difficult. Exploring the structure of the conditionally belief-free equilibrium payoffs set is the subject of our future research. However, the techniques that have been introduced by Fudenberg, Levine and Maskin (1994) or Kandori and Matsushima (1998) to analyze the structure of this set cannot be used in our setup. We need to develop similar but stronger techniques for this purpose.

6 Appendix

Proof of Proposition 3.1. Assume first that there is a pure strategy $\widehat{a}$ with $u(\widehat{a}) = v$, and consider the following strategy:

In period 1, each player i plays $\widehat{a}_i$. Each player i continues to play $\widehat{a}_i$ so long as the realized actions were $\widehat{a}$ in all previous periods. If at least one player did not play according to $\widehat{a}$, then each player i plays α_i^* until nature moves *Forget*. Once players forget, every player i starts playing $\widehat{a}_i$ until at least one player does not play according to $\widehat{a}$.

We need to check that this strategy is a Perfect Bayesian Equilibrium. For this we need to show that the strategy satisfies the no improvement property.

Define $\overline{v} = \max_{a_i} u_i(a_i, \widehat{a}_{-i})$ and suppose that up until period t no player has deviated. If player i plays according to $\widehat{a}_i$ at every period t, his utility will be v_i. If he deviates at period t and reverts to his initial strategy, however, the gain will be

$$(1-\delta)\overline{v} + \delta \left[\frac{(1-\delta)(1-p)}{1-\delta(1-p)} y_i + \frac{p}{1-\delta(1-p)} v_i \right] \tag{3}$$

Therefore, player i never deviates as long as the payoff given in eqn.3 is less than or equal to v_i. This is expressed by following inequality:

$$(1-\delta)\overline{v} + \delta \frac{(1-\delta)(1-p)}{1-\delta(1-p)} y_i \leq \left(1 - \frac{\delta p}{1-\delta(1-p)} \right) v_i \tag{4}$$

which is equivalent to

$$(1-\delta(1-p))\,\overline{v} + \delta(1-p) y_i \leq v_i \tag{5}$$

Eqn.5 can be rewritten as

$$\overline{v} - \delta(1-p)\overline{v} + \delta(1-p) y_i \leq v_i \tag{6}$$

or equivalently

$$\delta(1-p)\left[y_i - \overline{v}\right] \leq v_i - \overline{v} \tag{7}$$

We know that $y_i < v_i$ and $v_i \leq \overline{v}$. So, by multiplying both sides with negative one, we get

$$\delta(1-p)\left[\overline{v} - y_i\right] \geq \overline{v} - v_i \tag{8}$$

so, choose $x = \overline{\delta}(1-\overline{p})$ such that

$$\overline{\delta}(1-\overline{p})\left[\overline{v} - y_i\right] = \overline{v} - v_i \tag{9}$$

Note that for every $\delta > \overline{\delta}$ and every $p < \overline{p}$ eqn.8 and hence eqn.5 will be satisfied, meaning that player i has no incentive to deviate. It is important to see whether there exists $\overline{\delta}, \overline{p}$ with $\overline{p}, \overline{\delta} \in (0,1)$ that satisfies eqn.9: For that solve $\overline{\delta}$ as a function of $\overline{p}$:

$$\overline{\delta} = \left(\frac{1}{1-\overline{p}}\right)\left[\frac{\overline{v}-v_i}{\overline{v}-y_i}\right] \tag{10}$$

Note that we have $0 < \overline{\delta}$. We may have $\overline{\delta} < 1$ if and only if

$$\left(\frac{1}{1-\overline{p}}\right)\left[\frac{\overline{v}-v_i}{\overline{v}-y_i}\right] < 1 \tag{11}$$

Which is equivalent to

$$\overline{p} < 1 - \left[\frac{\overline{v}-v_i}{\overline{v}-y_i}\right] \tag{12}$$

Since $v_i > y_i$, we have $\frac{\overline{v}-v_i}{\overline{v}-y_i} < 1$. So, there exists some $\overline{p}$ and $\overline{\delta}$ with $0 < \overline{p}, \overline{\delta} < 1$ such that eqn.9 is satisfied.

On the other hand, if there is no pure strategy $\widehat{a}$ with $u(\widehat{a}) = v$, we use public randomizations $a(w)$. In this case if player i does not deviate his payoff is

$$(1-\delta)u_i(a(w)) + \delta v_i \geq (1-\delta)\underline{v} + \delta v_i \tag{13}$$

where $\underline{v} = \min_a u_i(\widehat{a})$. The realization of $a(w)$ might give less than v_i but it cannot be less than $\underline{v}$. Therefore, player i does not deviate if, and only if

$$(1-\delta)\overline{v} + \delta\left[\frac{(1-\delta)(1-p)}{1-\delta(1-p)}y_i + \frac{p}{1-\delta(1-p)}v_i\right] \leq \delta v_i + (1-\delta)\underline{v} \tag{14}$$

which is equivalent to

$$(1-\delta)\left[\overline{v}-\underline{v}\right] + \frac{\delta(1-\delta)(1-p)}{1-\delta(1-p)}y_i \leq \delta v_i\left[1 - \frac{p}{1-\delta(1-p)}\right] \tag{15}$$

$$(1-\delta)\left[\overline{v}-\underline{v}\right] + \frac{\delta(1-\delta)(1-p)}{1-\delta(1-p)}y_i \leq \frac{\delta(1-\delta)(1-p)}{1-\delta(1-p)}v_i \tag{16}$$

$$(1-\delta(1-p))\left[\overline{v}-\underline{v}\right] + \delta(1-p)y_i \leq \delta(1-p)v_i \tag{17}$$

$$\left[\overline{v}-\underline{v}\right] + \delta(1-p)\left[y_i - \overline{v} + \underline{v}\right] \leq \delta(1-p)v_i \tag{18}$$

$$\overline{v}-\underline{v} \leq \delta(1-p)\left[v_i - y_i + \overline{v} - \underline{v}\right] \tag{19}$$

Therefore, choose $x = \overline{\delta}(1-\overline{p})$ such that eqn.19 holds with equality. Note that for every $\delta > \overline{\delta}$ and every $p < \overline{p}$ eqn.19 and hence eqn.14 will be satisfied, meaning that player i has no incentive to deviate. It is important to see whether there exists $\overline{\delta}, \overline{p}$ with $\overline{p}, \overline{\delta} \in (0,1)$ that satisfies eqn.19 with equality: For that solve $\overline{\delta}$ as a function of $\overline{p}$:

$$\overline{\delta} = \left(\frac{1}{1-\overline{p}}\right)\left[\frac{\overline{v}-\underline{v}}{\overline{v}-\underline{v}+v_i-y_i}\right] \tag{20}$$

So that $0 < \overline{\delta}$. Moreover, $\overline{\delta} < 1$ if

$$\left(\frac{1}{1-\overline{p}}\right)\left[\frac{\overline{v}-\underline{v}}{\overline{v}-\underline{v}+v_i-y_i}\right] < 1 \tag{21}$$

implies

$$\overline{p} < 1 - \left[\frac{\overline{v}-\underline{v}}{\overline{v}-\underline{v}+v_i-y_i}\right] \tag{22}$$

Since $v_i > y_i$, we have $\frac{\overline{v}-\underline{v}}{\overline{v}-\underline{v}+v_i-y_i} < 1$. So, there exists some $\overline{p}$ and $\overline{\delta}$ with $\overline{p}, \overline{\delta} \in (0,1)$ such that eqn.19 is satisfied with equality.

Proof of Proposition 3.2. Assume first that there is a strategy $\widehat{a}$ (possibly public randomization) with $u(\widehat{a}) = v$, where $v \in V^*$. Choose $v' = (v'_1, ..., v'_n)$ in the interior of V^* such that $v_i > v'_i$ for all i. Since v' is the interior of V^* and V^* has full dimension, there exists $\epsilon > 0$ such that for each j, $(v'_1 + \epsilon, ..., v'_{j-1} + \epsilon, v'_j, v'_{j+1} + \epsilon, ..., v'_n + \epsilon) \in V^*$.

Suppose that $a(j) = (a_1(j), ..., a_n(j))$ be a joint strategy such that $u(a(j)) = (v'_1 + \epsilon, ..., v'_{j-1} + \epsilon, v'_j, v'_{j+1} + \epsilon, ..., v'_n + \epsilon)$. Moreover, assume that $w_i^j = u_i(m^j)$. Now, consider the following strategy:

Play begins in *phase I*. In phase I, play action profile $\widehat{a}$, where $u(\widehat{a}) = v$. Play remains in phase I so long as in each period either the realized action is $\widehat{a}$, or the realized action differs from $\widehat{a}$ in two or more components. If a single player j deviates from $\widehat{a}$, then play moves to phase II_j.

Phase II_j play m^j (minimax strategy) each period. Continue phase II_j for N periods so long as in each period either the realized action is m^j or the realized action differs from m^j in two or more components. Switch to phase III_j after N successive periods of phase II_j. If during phase II_j a single player i's action differs from m_i^j, begin phase II_i. (Note that this construction makes sense only if mixed actions are observable).[10]

Phase III_j Play $\widehat{a}(j)$, and continue to do so unless in some period a single player i fails to play $\widehat{a}_i(j)$. If a player i does deviate, begin phase II_i.

[10] Fudenberg and Maskin (1986) show that their proof can be replied even if mixed actions are not observable. We believe similar replication for our proof can easily be verified.

If nature takes the action *Forget*, then players move to phase I regardless of the phase that they were in. We need to consider the following three exhaustive cases:

Deviation in Phase I: If player i does not deviate in phase I, his payoff is $\frac{v_i}{1-\delta}$ (this is not average payoffs). If he deviates, however, his payoff is

$$\overline{v} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))} v_i + \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)} v_i' \tag{23}$$

Therefore, net gain from deviation is less than

$$\overline{v} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))} v_i + \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)} v_i' - \frac{v_i}{1-\delta} \tag{24}$$

Since $\frac{1}{1-\delta} = \frac{1}{1-\delta(1-p)} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))}$, eqn.23 is less than

$$\overline{v} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))} v_i + \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)} v_i' - \frac{v_i}{1-\delta(1-p)} - \frac{\delta p}{(1-\delta)(1-\delta(1-p))} v_i \tag{25}$$

which is also less than

$$\overline{v} - v_i' \left[\frac{1-\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)} \right] \tag{26}$$

In limit, eqn.24 is less than $\overline{v} - (N+1)v_i'$ and for $N \geq \frac{\overline{v}-v_i'}{v_i'}$, the net gain from deviation is negative.

CASE 2: Deviation in phase II

Subcase 1: Player i deviates in phase II_i: Then he will get at most zero for the period in which he deviates (because others are minimaxing him and his best reply gives zero) and then only lengthens his punishments, postponing the positive payoff of v_i'. So he will not deviate.

Subcase 2: Player i deviates in II_j. If player i does not deviate, he would get w_i^j (payoff of minimaxing player j, which might be negative) for $K < N$ periods and then gets $v_j'(i) = v_i' + \epsilon$. So if he deviates, the payoff of deviating will be

$$\frac{p}{(1-\delta)(1-\delta(1-p))} v_i + \frac{(1-p)(1-\delta^K(1-p)^K)}{1-\delta(1-p)} w_i + \frac{(1-p)^{K+1}\delta^K}{1-\delta(1-p)} (v_i' + \epsilon) \tag{27}$$

If player i deviates; his payoff will be same as eqn.23.

Therefore, the net gain from deviation in this case is

$$\overline{v} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))} v_i + \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)} v_i' - \frac{p}{(1-\delta)(1-\delta(1-p))} v_i - \frac{(1-p)(1-\delta^K(1-p)^K)}{1-\delta(1-p)} w_i - \frac{(1-p)^{K+1}\delta^K}{1-\delta(1-p)} (v_i' + \epsilon) \tag{28}$$

By multiplying the forth term by δ, the value of eqn.28 will increase. Following that, we can cancel the second and the forth terms. Fifth term can be rewritten as $-\left(\frac{1-\delta^n(1-p)^{n+1}}{1-\delta(1-p)}\right)w_i + \frac{p}{1-\delta(1-p)}w_i$ where first term converges to a finite number K as $\delta(1-p)$ converges to 1, and second term is always less than 1 (this is because $p < 1-\delta(1-p) \Leftrightarrow \delta(1-p) < 1-p \Leftrightarrow \delta < 1$).

Third and sixth terms of eqn.28 are no more than

$$\frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}v_i' - \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}v_i' - \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}\epsilon \tag{29}$$

The last term of eqn.29 converges to $-\infty$ as $\delta(1-p)$ converges to 1. Hence, the net gain in this case is negative as well. That is to say, there exists some $x \in (0,1)$ such that for all δ, p with $\delta(1-p) > x$ the net gain falls below zero. Hence player i has no incentive to deviate in Phase II_j.

CASE 3: Deviation in Phase III

Subcase 1: Player i deviates in III_i: In this case player i gets payoff as given in eqn.23. However, if he does not deviate, he gets the payoff of

$$\frac{p}{(1-\delta)(1-\delta(1-p))}v_i + \frac{1-p}{1-\delta(1-p)}v_i' \tag{30}$$

Therefore, net gain from deviation in this case is less than

Therefore, the net gain from deviation in this case is

$$\begin{aligned}&\overline{v} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))}v_i + \\ &\frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}v_i' - \frac{p}{(1-\delta)(1-\delta(1-p))}v_i - \frac{1-p}{1-\delta(1-p)}v_i'\end{aligned} \tag{31}$$

Multiply the forth term in eqn.31 with δ so that it cancels out with the second term, and net gain from deviation is still no more than the resulting argument. Besides, sum of the third term and the last term is less than or equal to

$$-(1-p)\left[\frac{1-\delta^N(1-p)^N}{1-\delta(1-p)}\right]v_i' \tag{32}$$

As $\delta(1-p)$ converges to 1, the term in eqn.32 converges to $-Nv_i'$. Therefore, net gain from deviation is no more than $\overline{v} - Nv_i'$ which is less than zero as long as $N > \frac{\max U}{v_i'}$.

Therefore, if we choose N such that $N > \max_{i \in n}\{\frac{\overline{v}-v_i'}{v_i'}, \frac{\overline{v}}{v_i'}\}$, then no player has incentive to deviate in any case.

Subcase 2: Player i deviates in III_j: If he deviates his payoff is same as in eqn.23. If he does not deviate, however, he gets

$$\frac{p}{(1-\delta)(1-\delta(1-p))}v_i + \frac{1-p}{1-\delta(1-p)}(v_i' + \epsilon) \tag{33}$$

Therefore, the net gain form deviation is less than

$$\begin{aligned} &\overline{v} + \frac{\delta p}{(1-\delta)(1-\delta(1-p))}v_i + \\ &\frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}v_i' - \frac{p}{(1-\delta)(1-\delta(1-p))}v_i - \frac{1-p}{1-\delta(1-p)}(v_i' + \epsilon) \end{aligned} \tag{34}$$

Multiply the forth term of eqn.34 with δ so that second and forth terms sum up to zero. In addition, sum of the third and the fifth terms are less than

$$\frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}v_i' - \frac{\delta^{N+1}(1-p)^{N+1}}{1-\delta(1-p)}v_i' - \left(\frac{1-p}{1-\delta(1-p)}\right)\epsilon \tag{35}$$

Note that the last term in eqn.35 converges to $-\infty$ as $\delta(1-p)$ converges to 1. Hence, there exists some $x^* \in (0,1)$ such that for all $\delta(1-p) > x^*$ deviation is not profitable.

It is crucial to note that N as defined previously depends on player i's payoff values. We can choose N such that $N > \max\{N_1, ..., N_n\}$ that provides enough incentive for every player not to deviate from their initial strategy.

This last subcase finishes the proof of Theorem 2.

Proof of Proposition 4.1. We prove a stronger claim: Let $s = (s_1, s_2)$ is a conditionally belief-free equilibrium strategy profile with full support over the regime sequence $\{\mathcal{A}^t\}$. Wlog, consider player 1, and let s_1^* be a strategy with support (not necessarily full support) over the regime sequence $\{\mathcal{A}^t\}$, i.e., for every t, and $h_1^t \in H_1^t$, $s_1^*(h_1^t) \in \Delta\mathcal{A}_1^t$. We claim that s_1^* is a conditionally belief-free sequential best reply to s_2, i.e., $s_1^*|_{h_1^t} \in B_1(s|_{h_2^t})$ for all t, $h_1^t \in H_1^t$ and $h_2^t \in \widetilde{h_1^t}$. Wlog, assume that s_1^* is a pure strategy. For each period t, and history h_1^t, the mixture $s_1(h_1^t)$ assigns positive probability to the pure action $s_1^*(h_1^t)$, simply because $s_1^*(h_1^t)$ is an element of $\mathcal{A}_1^t$ and $\mathcal{A}_1^t$ is the regime governing the strategy s in period t.

Therefore, for any t and h_1^t, define a continuation strategy $\hat{s}_1|_{h_1^t}$ to be the strategy which begins by playing the pure action $s_1^*(h_1^t)$ in period t and plays according to $s_1|_{h_1^t}$ after period t. By definition, the continuation strategy $\hat{s}_1|_{h_1^t} \in B_1(s|_{h_2^t})$ for all t, $h_1^t \in H_1$

and $h_2^t \in \widetilde{h_1^t}$. This is because $s_1|_{h_1^t}$ is a best response to $s_2|_{h_2^t}$ and $\hat{s}_1|_{h_1^t}$ differs from $s_1|_{h_1^t}$ only in period t in which $s_1(h_1^t)$ assigns positive probability to the pure action $\hat{s}_1(h_1^t)$.

We, now, construct a sequence of strategies for player 1, s_1^t, for $t = 0, 1, ...$ First, set $s_1^0 = s_1^*$. Inductively define s_1^t by $s_1^t(h_1^\tau) = s_1^{t-1}(h_1^\tau)$ if $\tau < t$ and $s_1|_{h_1^t} = \hat{s}_1|_{h_1^t}$. By construction, we have $s_1^t\big|_{h_1^t} = \hat{s}_1|_{h_1^t}$ which is a best reply to $s_2|_{h_2^t}$ for all t, $h_1^t \in H_1^t$, and $h_2^t \in \widetilde{h_1^t}$.

Now, by replacing the continuation strategy of $s_1^t\big|_{h_1^t}$ by $\hat{s}_1|_{h_1^{t+1}}$, we can derive $s_1^{t+1}\big|_{h_1^t}$. Then by construction, we must have $s_{1h_1^t}^{t+1} \in B_1(s_2|_{h_2^t})$, we can inductively conclude that

$$s_1^{t+k}|_{h_1^t} \in B_1(s|_{h_2^t}) \text{ for all } k \geq 0, h_1^t \in H_1^t \text{ and } \; h_2^t \in \widetilde{h_1^t}$$

By construction, for all $k \geq 0, s_1^{t+k}(h_1^t) = s_1^*(h_1^t)$ and thus for any fixed h_1^t, the sequence of continuation strategies $s_1^{t+k}\Big|_{h_1^t}$ converges to $s_1^*|_{h_1^t}$ as $k \to \infty$, history by history (in the product topology). Since discounted payoffs are continuous in the product topology, 6 implies that $s_1^*|_{h_1^t} \in B_1(s|_{h_2^t})$ for all t, $h_1^t \in H_1^t$ and $h_2^t \in \widetilde{h_1^t}$ as required

Proof of Proposition 4.2. Let $v = (v_i, v_{-i}) \in W$. We will show that player i has a strategy, s_i, which randomizes over the set $\mathcal{A}_i$, after each history, against which player $-i$'s maximum payoff is v_{-i} and that this payoff is achieved by any strategy which randomizes over $\mathcal{A}_{-i}$. Since the symmetric arguments imply the same conclusion with the roles reversed, these strategies form a *conditionally belief-free* strategy profile s with full support over the constant regime $\mathcal{A}$.

Construct the Markovian strategy s_i as follows. Start with the payoff vector $v = (v_i, v_{-i}) \in W$. Since W is strongly self-generating, for player $-i$, there exists some self-generating $\mathbf{F}$-collection $\overline{W}_{-i}$ and an $\mathbf{F}$-generated set X such that $v_{-i} \in X \in \overline{W}_{-i}$. Without loss of generality, we can take $X = \{v_{-i}\}$. Since, the $\mathbf{F}$-collection $\overline{W}_{-i}$ is self-generated, there must exists a mixture $\alpha_i^{v_{-i}} \in \Delta\mathcal{A}_i$ and a continuation function $w_{-i}^{v_{-i}}$ such that equation 4.2 holds for all $a_{-i} \in \mathcal{A}_{-i}$ i.e., $(\alpha_i^{v_{-i}}, w_{-i}^{v_{-i}})$ generates v_{-i}. Therefore, assign $s_i(1) = \alpha_i^{v_{-i}}$.

Since $\overline{W}_{-i}$ is self-generating $\mathbf{F}$-collection, the set $U^2 = \{w_{-i}^{v_{-i}}(a, y_i)\}_{a \in \mathbf{A}, y_i \in \Sigma}$ is a subset of some set in $\overline{W}_{-i}$. So, the "state" of the strategy in period 2 will be the continuation value for player $-i$, call it $u = w_{-i}^{v_{-i}}(a, y_i) \in U^2$, where (a, y_i) is player i's realization of last period private history. Again, since $\overline{W}_{-i}$ is self-generating and U^2 is a subset of some set in $\overline{W}_{-i}$, for each state $u \in U^2$, we know that there exists a mixture

$\alpha_i^u \in \mathcal{A}_i$ and a continuation function w_{-i}^u such that equation 4.2 holds, i.e., (α_i^u, w_{-i}^u) generates $u \in U^2$. Therefore, after any one-length history (a, y_i) and the associated state u, set $s_i(a, y_i) = \alpha_i^u$.

Remark that the set U^2 is an $\mathbf{F}$- generated and $\overline{W}_{-i}$ is self-generating $\mathbf{F}$-collection, i.e., the set $U^3 = \{w_{-i}^u(a, y_i)\}_{u \in U_1, a \in \mathbf{A}, y_i \in \Sigma}$ is a subset of some set in $\overline{W}_{-i}$, and for each $u, u' \in U^2$, $w_{-i}^u(., y_i) = w_{-i}^{u'}(., y_i)$ whenever $y_i = \mathbf{F}$. Therefore, when in period 2 the state u has realized and having played the action profile $a' \in \mathbf{A}$ and observed the private signal $y_i' \in \Sigma$ in stage 2, player i will transit to state $u' = w_{-i}^u(a', y_i') \in U^3$ in stage 3. The $\mathbf{F}$- generation guarantees that there is a unique state $u' \in U^3$ for each 2-length private history for player i, and whenever two 2-length private history have same memories, the states associated to each histories are same.

Continue inductively to construct the strategy profile s this way. We can ensure that for any t, k $(\leq t)$, player i, and for any two histories $h, h' \in \overline{h_i^t(k)}$ we have $s_i(h) = s_i(h')$. Moreover, it follows from Equation 4.2 and the *one-shot deviation* property that any such strategy profile $s = (s_i, s_{-i})$ is a *conditionally belief-free* strategy with full support over the regime $\mathcal{A}$, and achieves the payoff of v as required.

Proof of Proposition 4.3. Pick a conditionally belief-free equilibrium strategy profile s with full support over the constant regime $\mathcal{A}$, and let $w_i(h) = u_i(s|_h)$ for the continuation payoff to player i after history $h \in H$. Define W_i^s be the set of all possible continuation values in equilibrium s. Formally, $W_i^s = \{w_i(h) : h \in H\}$. We first show that $W^s = W_i^s \times W_{-i}^s$ is strongly self-generating. Since the conditionally belief free equilibrium strategy s is arbitrary, and the union of self generating sets is also self generating, this will lead us to $E = \cup_{s \in S} W^s$ is strongly self-generating as well, where S is the set of all conditionally belief-free equilibrium strategy profiles with full support over the constant regime $\mathcal{A}$. We prove each of these claims in turn.

Observe, first, that for a given conditionally belief-free equilibrium, $w_i(h)$ depends only on the *memory* of player $-i$'s history h_{-i}. To prove this point, first, fix a t-length history h_{-i} with memory k. Denote $m^{h_{-i}}$ as the memory of the history h_{-i}. Then, consider three t-length histories h_i, h_i' with memories less than k and h_i'' with memory more than k such that $(h_i, h_{-i}), (h_i', h_{-i})$ and $(h_i', h_{-i}) \in H$.[11]

[11] Restricting attention to histories h_i such that $(h_i, h_{-i}) \in H$ does not alter our analysis. Suppose that h_i's memory is less than h_{-i}'s memory and h_i's memory does not contradict with h_{-i}'s memory but $(h_i, h_{-i}) \notin H$. We can always find a history $\hat{h}_i$ that has same memory with h_i, so that $s_i|_{h_i} = s_i|_{\hat{h}_i}$, and $(\hat{h}_i, h_{-i}) \in H$. Therefore, $s_i|_{h_i} \in B_i(s_{-i}|_{h_{-i}})$ and $s_{-i}|_{h_{-i}} \in B_{-i}(s_i|_{h_i})$. On the other hand, if

We must have $w_i(h_i, h_{-i}) = w_i(h'_i, h_{-i})$. Suppose for a contradiction that $w_i(h_i, h_{-i}) > w_i(h'_i, h_{-i})$ implying that $s_i|_{h'_i}$ is not best reply to $s_{-i}|_{h_{-i}}$ contradicting that s is a conditionally belief-free strategy.

Now, consider h_i and h''_i. We claim that $w_i(h_i, h_{-i}) = w_i(h''_i, h_{-i})$. First, observe that $s_{-i}|_{h_{-i}}$ is best reply to $s_i|_{h''_i}$, and vice versa. Now suppose for a contradiction that $w_i(h_i, h_{-i}) > w_i(h''_i, h_{-i})$; in that case player i never plays $s_i|_{h''_i}$ against $s_{-i}|_{h_{-i}}$ after observing the history h''_i which contradicts to the assertion that s is conditionally belief-free equilibrium strategy. However, if we assume $w_i(h_i, h_{-i}) < w_i(h''_i, h_{-i})$ than in that case player i never plays $s_i|_{h_i}$ against $s_{-i}|_{h_{-i}}$ after observing the history h_i which contradicts, again, to the assertion s is an equilibrium strategy. Therefore, we can denote $w_i(h^t) = w_i(m^{h^t_{-i}})$

Now, fix some arbitrary period $t+1$ and history $h^t \in H$. Suppose that the mixed action $\alpha_{-i} := s_{-i}(h^t_{-i}) \in \Delta\mathcal{A}_{-i}$ is played by i's opponent after h^t.Let $a^*_i \in \mathcal{A}_i$ be arbitrary. Then, there is a best reply continuation strategy $\widehat{s}_i$ for player i which plays a^*_i at period $t+1$ after history h^t_i. That is, $\widehat{s}_i$ is a best reply continuation strategy to $s_{-i}|_{h^t_{-i}}$. Moreover, $u_i(\widehat{s}_i, s_{-i}|_{h^t_{-i}}) = u_i(s|_{h^t}) = w_i(h^t)$.

We claim that the payoff vector $w_i(h^t)$ can be generated by W^s_i. For that purpose, we, first, need to show that there exists a continuation payoff function w_i which depends only on h^t_{-i}, a^{t+1} and y^{t+1}_{-i} i.e., the memory of h^{t+1}_{-i}, $m^{h^{t+1}_{-i}} = m^{(h^t_{-i}, a^{t+1}, y^{t+1}_{-i})}$.

First consider the case where $y^{t+1}_{-i} = \mathbf{F}$. For any $a^{t+1} \in \mathbf{A}$, and $y_i \in \Sigma$ we claim that $w_i(h^{t+1}_{-i})$ yields the same value: Take any $a, a' \in \mathbf{A}$, and suppose first that $y^{t+1}_i = \mathbf{F}$. Then we know that any continuation strategy of player i after observing $(a, \mathbf{F})$ at stage $t+1$ is also a continuation strategy after observing $(a', \mathbf{F})$ and are best reply to $s_{-i}|_{h^{t+1}_{-i}}$ and vice versa, where $h^{t+1}_{-i} = (h^t_{-i}, a, \mathbf{F})$, since both players have memory zero. Therefore, these continuation strategies of player i will automatically achieve the same value. Now, consider the case where $y^{t+1}_i = \mathbf{R}$. Since $y^{t+1}_{-i} = \mathbf{F}$, the continuation strategy $s_{-i}|_{h^{t+1}_{-i}}$ is best reply to $s_i|_{(h^t_i, a, \mathbf{R})}$ and $s_i|_{(h^t_i, a', \mathbf{R})}$. Although player i's best reply continuation strategies after observing $(a, \mathbf{R})$ and $(a', \mathbf{R})$ at stage $t+1$ may be different, they have to yield same continuation values i.e., both are best reply to $s_{-i}|_{h^{t+1}_{-i}}$. Suppose not, i.e., $w_i(h^t_i; a, \mathbf{R}) = u_i(s_i|_{(h^t_i, a, \mathbf{R})}, s_{-i}|_{h^{t+1}_{-i}}) > w_i(h^t_i; a', \mathbf{R}) = u_i(s_i|_{(h^t_i, a', \mathbf{R})}, s_{-i}|_{h^{t+1}_{-i}})$. Therefore, player i deviates after history $(h^t_i, a', \mathbf{R})$ and plays $s_i|_{(h^t_i, a, \mathbf{R})}$, which contra-

h_i's memory is more than h_{-i}'s memory, and h_{-i}'s memory does not contradict with h_i's memory but $(h_i, h_{-i}) \notin H$, we can always find a history $\hat{h}_{-i}$ that has same memory with h_{-i}, so that $s_{-i}|_{h_{-i}} = s_{-i}|_{\hat{h}_{-i}}$, and $(h_i, \hat{h}_{-i}) \in H$. Hence, $s_i|_{h_i} \in B_i(s_{-i}|_{h_{-i}})$ and $s_{-i}|_{h_{-i}} \in B_{-i}(s_i|_{h_i})$.

dicts with the fact that s is a conditionally belief free equilibrium. The same arguments for the reversed inequality shows that best reply continuation strategies must yield the same value. Therefore, we can conclude that if $y_{-i}^{t+1} = \mathbf{F}$, it really does not matter what action-signal profile player i has observed at stage $t+1$.

Now, assume that $y_{-i}^{t+1} = \mathbf{R}$ and $a^{t+1} = a = (a_i^*, a_{-i}) \in \mathbf{A}$. Then we need to show that whether player i observes the signal $\mathbf{F}$ or $\mathbf{R}$, his continuation strategies must give same values.[12] First observe that $s_{-i}|_{h_{-i}^{t+1}}$ is best reply to both $s_i|_{(h_i^t,a,\mathbf{R})}$ and $s_i|_{(h_i^t,a,\mathbf{F})}$ and vice versa. Although these best reply continuation strategies might be different, they have to achieve same values: Suppose for a contradiction that $w_i(h_i^t; a, \mathbf{R}) = u_i(s_i|_{(h_i^t,a,\mathbf{R})}, s_{-i}|_{h_{-i}^{t+1}}) > w_i(h_i^t; a, \mathbf{F}) = u_i(s_i|_{(h_i^t,a,\mathbf{F})}, s_{-i}|_{h_{-i}^{t+1}})$ which contradicts that s is a conditionally belief free equilibrium strategy. Same arguments in the case of reversed inequality shows that player i's continuation strategies must attain same value in both cases.

In this way we can view $w_i(.)$ as a continuation payoff function, which depends only on the memory of h_{-i}^{t+1}. Moreover, since $m^{h_{-i}^{t+1}} = (m^{h_{-i}^t} \cup \{a\}, t+1)$ if $y_{-i}^{t+1} = \mathbf{R}$, and $m^{h_{-i}^{t+1}} = (\varnothing, t+1)$ if $y_{-i}^{t+1} = \mathbf{F}$, we can write i's continuation payoff function, after history h_i^{t+1}, as $w_i(m^{h_{-i}^t}; ., .)$ taking values in W_i with the property that $w_i(m^{h_{-i}^t}; a, y_{-i}) = w_i(m^{h_{-i}^t}; a', y_{-i})$ for all $a', a \in \mathbf{A}$ whenever $y_{-i} = \mathbf{F}$.

Now, we need to show that this continuation function will generate $w_i(h^t)$ along with the mixture $\alpha_{-i} = s_{-i}(h_{-i}^t)$: The payoff to i from using $\widehat{s}_i$ against $s_{-i}(h_{-i}^t)$ can thus be written

$$(1-\delta)u_i(a_i^*, \alpha_{-i}) + \delta E\left(w_i(m^{h_{-i}^t}; ., .) \middle| a_i^*, \alpha_{-i}, y_{-i}\right)$$

where

$$E\left(w_i(m^{h_{-i}^t}; ., .) \middle| a_i^*, \alpha_{-i}, y_{-i}\right) = \sum_{a_{-i} \in \mathbf{A}_{-i}} \sum_{y_{-i} \in \Sigma} \alpha_{-i}(a_{-i}) \pi(y_{-i}) w_i(m^{h_{-i}^t}; a_i^* a_{-i}, y_{-i})$$

Since $\widehat{s}_i$ is a best reply against $s_{-i}|_{h_{-i}^t}$, 6 is equal to $w_i(h^t)$. Moreover, since a_i^* was an arbitrary element of $\mathcal{A}_i$, this equality holds for all $a_i \in \mathcal{A}_i$. Finally, since s is conditionally belief-free strategy bounded by constant regime $\mathcal{A}$, player i cannot achieve a greater continuation value with a strategy that begins with some action $a_i \notin \mathcal{A}_i$. Thus, $w_i(h^t)$ must be greater than or equal to expression 6 for actions $a_i \notin \mathcal{A}_i$

[12] Given that player $-i$ observes the signal (a, R), we do not need to consider the case where player i observes some $a \neq a'$, because such history will never be in H.

Therefore, the mixture $s_{-i}(h^t_{-i})$ enforces $w_i(m^{h^t_{-i}};.,.)$ and generates $w_i(h^t) \in W^s_i$. Since h^t was arbitrary, every element of W^s_i can be so generated. Applying the same arguments for player $-i$ shows that the set $W^s = W^s_i \times W^s_{-i}$ is self-generating.

To show, W is strongly self generating, we need to show that for any i, there exists a self-generating $\mathbf{F}$-collection $\overline{W}^s_i$ such that $W^s_i \subseteq B^*(\overline{W}^s_i)$. So, for given conditionally belief-free strategy s and for each player i, let $O^0 = \{w_i(\varnothing)\}$, and for $t > 0$, $O^t_i = \{w_i(m^{h^{t-1}_{-i}}; a, y_{-i}) \mid h^t_{-i} \in H^t_{-i},\ a \in \mathbf{A}$ and $y_{-i} \in \Sigma_{-i}\}$. Then, for each player i, set $\overline{W}^s_i = \{O^t_i\}_t$.

Note that for each $t \geq 0$, O^t_i is generated by the set O^{t+1}_i i.e., for every real number $v \in O^t_i$, there exists a mixture $s_{-i}(h^{t-1}_{-i}, a^t, y^t_{-i}) \in \Delta\mathcal{A}_{-i}$ that enforces the continuation payoff function $w_i(m^{(h^{t-1}_{-i}, a^t, y^t_{-i})};.,.)$, which is in O^{t+1}_i for every $a^{t+1} \in \mathbf{A}, y^{t+1}_{-i} \in \Sigma_{-i}$, and generates v. Moreover, each O^t_i is $\mathbf{F}$-generated because, for any $a, a' \in \mathbf{A}$, $w_i(m^{(h^{t-1}_{-i}, a^t, y^t_{-i})};\ a, y_{-i}) = w_i(m^{(h^{t-1}_{-i}, a^t, y^t_{-i})};\ a', y_{-i})$ whenever $y_{-i} = \mathbf{F}$, because both histories yield same memories $(\varnothing, t+1)$. Therefore, $\overline{W}^s_i$ is a self-generating $\mathbf{F}$-collection. Moreover, we have $W^s_i = B^*(\overline{W}^s_i)$ for each i. Hence, W^s is strongly self-generating set.

Since E is the union of all continuation values occurring along histories of all conditionally belief-free equilibria, i.e., $E = \cup_{s\in S} W^s$ where S is the set of all conditionally belief-free equilibrium strategy profiles bounded by constant regime $\mathcal{A}$, it is the union of strongly self-generating sets and is therefore strongly self-generating; for each player i, define $\overline{W}_i = \{\overline{W}^s_i\}_{s\in S}$. It is easy to show that this collection is a self-generating $\mathbf{F}$-collection. Moreover, for each player i, we have $E_i = B^*(\overline{W}_i)$. Hence, E is strongly self-generating set.

Proof of Proposition 4.4. Let $v = (v_i, v_{-i}) \in W$, we will show that player i has a strategy, s_i, with full support according to μ such that player $-i$'s maximum payoff is v_{-i} and this payoff is achieved by any strategy for player $-i$ with full support according to μ. Since the symmetric arguments imply the same conclusion with the roles reversed, these strategies form a *conditionally belief-free* strategy profile s.

Construct the Markovian strategy s_i as follows. Start with the payoff vector $v_{-i} \in W_{-i}$. Since W is strongly self-generating, for player $-i$, there exists some self-generating $\mathbf{F}$-collection $\overline{W}_{-i}$ and an $\mathbf{F}$-generated set X such that $v_{-i} \in X \in \overline{W}_{-i}$. Without loss of generality, we can take $X = \{v_{-i}\}$. Since, the $\mathbf{F}$-collection $\overline{W}_{-i}$ is self-generated, there must exists a mixture $\alpha_i^{\mathcal{A},v_{-i}} \in \Delta\mathcal{A}_i$ and a continuation payoff function $w_{-i}^{\mathcal{A},v_{-i}}$ for each $\mathcal{A} \in \mathcal{P}(\mathbf{A})$ such that equation 1 holds for all $a_{-i}(\mathcal{A}) \in \mathcal{A}_{-i}$ for each $\mathcal{A} \in \mathcal{P}(\mathbf{A})$, i.e.,

$(\{\alpha_i^{\mathcal{A},v_{-i}}\}, \{w_{-i}^{\mathcal{A},v_{-i}}\})$ generates v_{-i}. Therefore, assign $s_i(1) = \alpha_i^{\mathcal{A}}$ after each associated public randomization $\mathcal{A}$.

Since $\overline{W}_{-i}$ is self-generating **F-** collection, the set $U^2 = \{w_{-i}^{\mathcal{A},v_{-i}}(a, y_i)\}_{\mathcal{A}\in\mathcal{P}(\mathbf{A}),a\in\mathbf{A},y_i\in\Sigma}$ is a subset of some set in $\overline{W}_{-i}$. So, the "state" of the strategy in period 2 will be the continuation value for player $-i$, call it $u = w_{-i}^{\mathcal{A},v_{-i}}(a, y_i) \in U^2$, where (a, y_i) is player i's realization of last period private history and $\mathcal{A}$ is the realization of the public randomization. Again, since $\overline{W}_{-i}$ is self-generating and U^2 is a subset of some set in $\overline{W}_{-i}$, for each state $u \in U^2$, we know that there exists a mixture $\alpha_i^{\mathcal{A},u} \in \Delta\ \mathcal{A}_i$ and a continuation function $w_{-i}^{\mathcal{A},u}$ for each $\mathcal{A}$ such that equation 1 holds, i.e., $(\{\alpha_i^{\mathcal{A},u}\}, \{w_{-i}^{\mathcal{A},u}\})$ generates $u \in U^2$. Therefore, after any one-length history $(a, y_i, \mathbf{z})$ and the associated state $u \in U^2$, set $s_i(a, y_i, \mathbf{z} = \mathcal{A}) = \alpha_i^{\mathcal{A},u}$.

Remark that the set U^2 is an **F-** generated and $\overline{W}_{-i}$ is self-generating **F**-collection, i.e., the set $U^3 = \{w_{-i}^{\mathcal{A},u}(a, y_i)\}_{\substack{\mathcal{A}\in\mathcal{P}(\mathbf{A}),u\in U^2\\ a\in\mathbf{A},y_i\in\Sigma}}$ is a subset of some set in $\overline{W}_{-i}$, and for each $u, u' \in U^2$ and $\mathcal{A} \in \mathcal{P}(\mathbf{A})$, $w_{-i}^{\mathcal{A},u}(., y_i) = w_{-i}^{\mathcal{A},u'}(., y_i)$ whenever $y_i = \mathbf{F}$. Moreover, when in period 2 the state u has realized and having played the action profile $a' \in \mathbf{A}$ and observed the private signal $y'_i \in \Sigma$ and public signal $\mathcal{A}$ in stage 2, player i will transit to state $u' = w_{-i}^{\mathcal{A},u}(a', y'_i) \in U^3$ in stage 3. The **F-** generation guarantees that there is a unique state $u' \in U^3$ for each 2-length private history for player i, and whenever two 2-length private history have same memories, the states associated to each histories are same.

Continue inductively to construct the strategy profile s in this way, we can ensure that for any t, $k(< t)$, player i, and for any two histories $h, h' \in \overline{h_i^t(k)}$ we have $s_i(h) = s_i(h')$. Moreover, it follows from Equation 1 and the *one-shot deviation* property that when the public randomization is *i.i.d* with distribution μ, any such strategy profile $s = (s_i, s_{-i})$is *conditionally belief-free* with full support according to μ, and achieves the payoff of v as required.

Proof of Proposition 4.5. Pick a conditionally belief-free equilibrium strategy profile s with full support according to μ, and let $w_i(h) = u_i(s|_h)$ for the continuation payoff to player i after history $h \in H$. Define W_i^s be the set of all possible continuation values in the equilibrium s. Formally, $W_i^s = \{w_i(h) : h \in H\}$. We first show that $W^s = W_i^s \times W_{-i}^s$ is strongly self-generating. Since the conditionally belief free equilibrium strategy s is arbitrary, and the union of self generating sets is also self generating, this will lead us to $E = \cup_{s\in S} W^s$ is strongly self-generating as well, where S is the set of all conditionally belief-free equilibrium strategy profiles with full support according to μ. We prove each of these claims in turn.

Observe, first, that for a given conditionally belief-free equilibrium, $w_i(h)$ depends only on the *memory* of player $-i$'s history h_{-i}. To prove this point, first, fix a t-length history h_{-i} with memory k. Denote $m^{h_{-i}}$ as the memory of the history h_{-i}. Then, consider three t-length histories h_i, h_i' with memories less than k and h_i'' with memory more than k such that $(h_i, h_{-i}), (h_i', h_{-i})$ and $(h_i', h_{-i}) \in H$.[13]

We must have $w_i(h_i, h_{-i}) = w_i(h_i', h_{-i})$. Suppose for a contradiction that $w_i(h_i, h_{-i}) > w_i(h_i', h_{-i})$ implying that $s_i|_{h_i'}$ is not best reply to $s_{-i}|_{h_{-i}}$ contradicting that s is a conditionally belief-free strategy.

Now, consider h_i and h_i''. We claim that $w_i(h_i, h_{-i}) = w_i(h_i'', h_{-i})$. First, observe that $s_{-i}|_{h_{-i}}$ is best reply to $s_i|_{h_i''}$, and vice versa. Now suppose for a contradiction that $w_i(h_i, h_{-i}) > w_i(h_i'', h_{-i})$; in that case player i never plays $s_i|_{h_i''}$ against $s_{-i}|_{h_{-i}}$ after observing the history h_i'' which contradicts to the assertion that s is conditionally belief-free equilibrium strategy.

However, if we assume $w_i(h_i, h_{-i}) < w_i(h_i'', h_{-i})$ than in that case player i never plays $s_i|_{h_i}$ against $s_{-i}|_{h_{-i}}$ after observing the history h_i which contradicts, again, to the assertion s is an equilibrium strategy. Therefore, we can denote $w_i(h^t) = w_i(m^{h^t_{-i}})$.

Now, fix some arbitrary period $t+1$ and history $h^t \in H$. Because s conforms to the public randomization, for each $\mathbf{z}^t = \mathcal{A} \in \mathcal{P}(A)$, the mixed action $\alpha_{-i} := s_{-i}(h^t_{-i}) \in \Delta\mathcal{A}_{-i}(\mathbf{z}^t)$ played by i's opponent after h^t where $\mathbf{z}^t$ is the realization of the public randomization in h^t. Let $a_i^*(\mathcal{A}) \in \mathcal{A}_i(\mathbf{z}^t)$. Since, $\mathcal{A}_i = \mathcal{A}_i^t(\mathbf{z}^t = \mathcal{A})$ for each $\mathcal{A}$, there is a best reply continuation strategy $\widehat{s}_i$ for player i which plays $a_i^*(\mathcal{A})$ at period $t+1$ after history h_i^t. That is, $\widehat{s}_i$ is a best reply continuation strategy to $s_{-i}|_{h^t_{-i}}$ for every $\mathcal{A} \in \mathcal{P}(A)$. Moreover, $u_i(\widehat{s}_i, s_{-i}|_{h^t_{-i}}) = u_i(s|_{h^t}) = w_i(h^t)$.

We claim that the payoff vector $w_i(h^t)$ can be generated by W_i^s. For that purpose, we, first, need to show that there exists a continuation payoff function w_i which depends only on h^t_{-i}, a^{t+1} and y^{t+1}_{-i} i.e., the memory of h^{t+1}_{-i}, $m^{h^{t+1}_{-i}} = m^{(h^t_{-i}, a^{t+1}, y^{t+1}_{-i})}$.

[13] Restricting attention to histories h_i such that $(h_i, h_{-i}) \in H$ does not alter our analysis. Suppose that h_i's memory is less than h_{-i}'s memory and h_i's memory does not contradict with h_{-i}'s memory but $(h_i, h_{-i}) \notin H$. We can always find a history $\hat{h}_i$ that has same memory with h_i, so that $s_i|_{h_i} = s_i|_{\hat{h}_i}$, and $(\hat{h}_i, h_{-i}) \in H$. Therefore, $s_i|_{h_i} \in B_i(s_{-i}|_{h_{-i}})$ and $s_{-i}|_{h_{-i}} \in B_{-i}(s_i|_{h_i})$. On the other hand, if h_i's memory is more than h_{-i}'s memory, and h_{-i}'s memory does not contradict with h_i's memory but $(h_i, h_{-i}) \notin H$, we can always find a history $\hat{h}_{-i}$ that has same memory with h_{-i}, so that $s_{-i}|_{h_{-i}} = s_{-i}|_{\hat{h}_{-i}}$, and $(h_i, \hat{h}_{-i}) \in H$. Hence, $s_i|_{h_i} \in B_i(s_{-i}|_{h_{-i}})$ and $s_{-i}|_{h_{-i}} \in B_{-i}(s_i|_{h_i})$.

First consider the case where $y_{-i}^{t+1} = \mathbf{F}$. For any $a^{t+1} \in \mathbf{A}$, and $y_i \in \Sigma$ we claim that $w_i(h_{-i}^{t+1})$ yields the same value: Take any two $a(\mathcal{A})$, $a'(\mathcal{A}) \in \mathbf{A}$, and suppose first that $y_i^{t+1} = \mathbf{F}$. Then, given that $\mathcal{A} \in \mathcal{P}(A)$ is publicly observed at stage $t+1$, we know that any continuation strategy after observing $(a(\mathcal{A}), \mathbf{F})$ at stage $t+1$ is also a continuation strategy after observing $(a'(\mathcal{A}), \mathbf{F})$ and are best reply to $s_{-i}|_{h_{-i}^{t+1}}$ and vice versa, where $h_{-i}^{t+1} = (h_{-i}^t, \mathcal{A}, a(\mathcal{A}), \mathbf{F})$, since both players have memory zero. Therefore, these continuation strategies of player i will automatically achieve the same value. Now, consider the case where $y_i^{t+1} = \mathbf{R}$. Since $y_{-i}^{t+1} = \mathbf{F}$, the continuation strategy $s_{-i}|_{h_{-i}^{t+1}}$ is best reply to $s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{R})}$ and $s_i|_{(h_i^t, \mathcal{A}, a'(\mathcal{A}), \mathbf{R})}$. Although player i's best reply continuation strategies after observing $(\mathcal{A}, a(\mathcal{A}), R)$ and $(\mathcal{A}, a'(\mathcal{A}), R)$ at stage $t+1$ may be different, they have to yield same continuation values. Suppose not, i.e., $w_i(h_i^t; \mathcal{A}, a(\mathcal{A}), R) = u_i(s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{R})}, s_{-i}|_{h_{-i}^{t+1}}) > w_i(h_i^t; \mathcal{A}, a'(\mathcal{A}), \mathbf{R}) = u_i(s_i|_{(h_i^t, \mathcal{A}, a'(\mathcal{A}), \mathbf{R})}, s_{-i}|_{h_{-i}^{t+1}})$. Therefore, player i deviates after history $(h_i^t, \mathcal{A}, a'(\mathcal{A}), \mathbf{R})$ and plays $s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{R})}$, which contradicts with the fact that s is a conditionally belief free equilibrium. The same arguments for the reversed inequality shows that best reply continuation strategies must yield the same value. Therefore, we can conclude that if $y_{-i}^{t+1} = \mathbf{F}$, it really does not matter what action-signal profile player i has observed at stage $t+1$.

Now, assume that $y_{-i}^{t+1} = \mathbf{R}$ and $a^{t+1} = a(\mathcal{A}) = (a_i^*(\mathcal{A}), a_{-i}(\mathcal{A})) \in \mathbf{A}$. Then we need to show that whether player i observes the signal $\mathbf{F}$ or $\mathbf{R}$, his continuation strategies must give same values.[14] First observe that $s_{-i}|_{h_{-i}^{t+1}}$ is best reply to both $s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{R})}$ and $s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{F})}$ and vice versa. Although these best reply continuation strategies might be different, they have to achieve same values: Suppose for a contradiction that $w_i(h_i^t; \mathcal{A}, a(\mathcal{A}), \mathbf{R}) = u_i(s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{R})}, s_{-i}|_{h_{-i}^{t+1}}) > w_i(h_i^t; \mathcal{A}, a(\mathcal{A}), \mathbf{F}) = u_i(s_i|_{(h_i^t, \mathcal{A}, a(\mathcal{A}), \mathbf{F})}, s_{-i}|_{h_{-i}^{t+1}})$ which contradicts that s is a conditionally belief free equilibrium strategy. Same arguments in the case of reversed inequality shows that player i's continuation strategies must attain same value in both cases.

In this way we can view $w_i(.)$ as a continuation value function, which depends only on the length and the memory of h_{-i}^{t+1}. Moreover, since $m^{h_{-i}^{t+1}} = (m^{h_{-i}^t} \cup \{a, \mathcal{A}\}, t+1)$ if $y_{-i}^{t+1} = \mathbf{R}$, and $m^{h_{-i}^{t+1}} = (\{\mathbf{z}^\tau\}_{\tau=1}^{\tau=t}, \mathcal{A}, t+1)$ if $y_{-i}^{t+1} = \mathbf{F}$, we can write i's continuation payoff function, after history h_i^{t+1}, as $w_i^{\mathcal{A}}(m^{h_{-i}^t}; ., .)$ taking values in W_i^s with the property that for any $\mathcal{A}$, $w_i^{\mathcal{A}}(m^{h_{-i}^t}; a(\mathcal{A}), y_{-i}) = w_i^{\mathcal{A}}(m^{h_{-i}^t}; a'(\mathcal{A}), y_{-i})$ for all $a(\mathcal{A}), a'(\mathcal{A}) \in \mathbf{A}$ whenever $y_{-i} = \mathbf{F}$.

[14] Given that player $-i$ observes the signal $(a(\mathcal{A}), \mathbf{R})$, we do not need to consider the case where player i observes some $a(\mathcal{A}) \neq a'(\mathcal{A})$, because such history will never be in H.

The payoff to i from using $\widehat{s}_i$ against $s_{-i}(h^t_{-i})$ when the realized regime in period $t+1$ is $\mathcal{A}$ can thus be written

$$(1-\delta)u_i(a_i^*(\mathcal{A}),\alpha_{-i}^{\mathcal{A}}) + \delta E(w_i^{\mathcal{A}}(m^{h^t_{-i}};.,.)\Big|\, a_i(\mathcal{A},\alpha_{-i}^{\mathcal{A}},y_{-i}) \tag{36}$$

where

$$E(w_i^{\mathcal{A}}(m^{h^t_{-i}};.,.)\Big|\, a_i(\mathcal{A}),\alpha_{-i}^{\mathcal{A}},y_{-i}) = \sum_{a_{-i}\in\mathbf{A}_{-i}}\sum_{y_{-i}\in\Sigma}\alpha_{-i}^{\mathcal{A}}(a_{-i})\pi(y_{-i})w_i^{\mathcal{A}}(m^{h^t_{-i}};a_i^*(\mathcal{A}),a_{-i},y_{-i}) \tag{37}$$

and the expected payoff before the realization of the regime in period $t+1$ is the expected value of this expression with respect to μ, i.e., the right hand side of equation 1.

Since $\widehat{s}_i$ is a best reply against $s_{-i}|_{h_{-i}}$, 36 is equal to $w_i(h^t)$. Moreover, since we selected i's action from each $\mathcal{A}_i(\mathbf{z}^{t+1}=\mathcal{A})$ arbitrarily, this equality holds for any selection $\{a_i^*(\mathcal{A})\}_{\mathcal{A}\in\mathcal{P}(A)}$. Finally, since s is conditionally belief-free, player i cannot achieve a greater continuation value with a strategy that does not conform to the regime. Thus, $w_i(h^t)$ must be greater than or equal to expression 36 when $a_i^*(\mathcal{A})\notin\mathcal{A}_i(\mathcal{A}_i^t)$ for some $\mathcal{A}$.

Therefore, the mixture sequence $\{s_{-i}(h^t_{-i},\mathcal{A})\}$ enforces the continuation payoff sequence $\{w_i^{\mathcal{A}}(m^{h^t_{-i}};.,.)\}$ for each $\mathcal{A}$ and generates $w_i(h^t)\in W_i^s$. Since h^t was arbitrary, every element of W_i^s can be so generated. Applying the same arguments for player $-i$ shows that the set $W^s=W_i^s\times W_{-i}^s$ is self-generating.

To show, W is strongly self generating, we need to show that for any i, there exists a self-generating $\mathbf{F}$-collection $\overline{W}_i^s$ such that $W_i^s\subseteq B^*(\overline{W}_i^s)$. So, for given conditionally belief-free strategy s with full support according to μ, and for each player i, let $O^0=\{w_i(\varnothing)\}$, and for $t>0$, $O_i^t=\{w_i^{\mathcal{A}}(m^{h^{t-1}_{-i}};a,y_{-i})\mid \mathcal{A}\in\mathcal{P}(A),\ h^t_{-i}\in H^t_{-i},\ a\in\mathbf{A}$ and $y_{-i}\in\Sigma_{-i}\}$. Then, for each player i, set $\overline{W}_i^s=\{O_i^t\}_t$.

Note that for each $t\geq 0$, O_i^t is generated by the set O_i^{t+1} i.e., for every real number $v\in O_i^t$, there exists a mixture $s_{-i}(h^{t-1}_{-i},a^t,y^t_{-i},\mathbf{z}^t)\in\Delta\mathcal{A}_{-i}(\mathbf{z}^t)$ that enforces the continuation payoff function $w_i(m^{(h^{t-1}_{-i},a^t,y^t_{-i},\mathbf{z}^t)};.,.)$, which is in O_i^{t+1} for every $a^{t+1}\in\mathbf{A},y^{t+1}_{-i}\in\Sigma_{-i}$, and generates v. Moreover, each O_i^t is $\mathbf{F}$-generated, because for any $a,a'\in\mathbf{A}$, $w_i(m^{(h^{t-1}_{-i},a^t,y^t_{-i},\mathbf{z}^t)};\ a,y_{-i}) = w_i(m^{(h^{t-1}_{-i},a^t,y^t_{-i},\mathbf{z}^t)};\ a',y_{-i})$ whenever $y_{-i}=\mathbf{F}$, because both histories yield same memories $(\varnothing,t+1)$. Therefore, $\overline{W}_i^s$ is a self-generating $\mathbf{F}$-collection. Moreover, we have $W_i^s=B^*(\overline{W}_i^s)$ for each i. Hence, W^s is a strongly self-generating set.

Since E is the union of all continuation values occurring along histories of all conditionally belief-free equilibria with full support according to μ, i.e., $E = \cup_{s \in S} W^s$ where S is the set of all conditionally belief-free equilibrium strategy profiles with full support according to μ, it is the union of strongly self-generating sets and is therefore strongly self-generating; for each player i, define $\overline{W}_i = \{\overline{W}_i^s\}_{s \in S}$. It is easy to show that this collection is a self-generating $\mathbf{F}$-collection. Moreover, for each player i, we have $E_i = B^*(\overline{W}_i)$. Hence, $E = E_1 \times E_2$ is a strongly self-generating set.

REFERENCES

Abreu, D., Pearce., D, and Stacchetti, E., " Optimal cartel monitoring with imperfect information", *Journal of Economic Theory* (1986) 39, 251-269.

Abreu, D., Pearce., D, and Stacchetti, E., "Toward a theory of discounted repeated games with imperfect monitoring", *Econometrica* (1990) 58, 1041-1063.

Aumann, R., Sorin, S., "Cooperation and bounded recall", *Games and Economic Behavior* (1989) 1, 5-39.

Cole, H., L., Kocherlakota, N.,R., "Finite memory and imperfect monitoring", *Games and Economic Behavior* (2005) 53, 59-72.

Ely, J., C., "A robust folk theorem for the Prisoner's dilemma", *Journal of Economic Theory* (2002) 102, 84-105.

Ely, J., C., Horner, J., and Olszewski, W., "Belief-free equilibria in repeated games", *Econometrica* (2005) 73, 377-415.

Friedman, J., "A non-cooperative equilibrium for supergames", *Review of Economic Studies* (1971)38, 1-12.

Fudenberg, D., Levine, D., and Maskin, E., "The folk theorem with imperfect public information", *Econometrica*, (1994) 62, 997-1039.

Fudenberg, D., and Maskin, E., "The folk theorem in repeated games with discounting and incomplete information", *Econometrica* (1986) 54, 533-554.

Fudenberg, D., and Tirole, J., (1991): Game Theory. MIT Press, Cambridge.

Kandori, M., and Matsushima, H., "Private observation, communication and collusion", *Econometrica*, (1998), 66, 627-652.

Kreps, D., Wilson, R., "Sequential equilibria", *Econometrica,* (1982) 50, 863 –894.

Lehrer, E., "Repeated games with stationary bounded recall strategies", *Journal of Economic Theory,* (1998) 46, 130-144.

Piccione, M., "The repeated prisoners' dilemma with imperfect private monitoring", *Journal of Economic Theory*, (2002) 102, 70-83.

Piccione, M., Rubinstein, A., "On the interpretation of decision problems with imperfect recall", *Games and Economic Behavior*, (1997) 20, 3-24.