Commitment games*

Ludovic Renou †

University of Leicester

February 19, 2008

Abstract

This paper explores how the ability to commit in games affect equilibrium payoffs. More precisely, we consider two-stage games, called *commitment games*, in which players can commit to some of their strategies in the first stage, and play the game induced by their commitment in the second stage. We completely characterize equilibrium payoffs of commitment games. Among others, we show that the power to commit in finitely repeated games as, for instance, finitely repeated prisoner's dilemma games, can lead to efficiency even though the constituent game does not satisfy the assumptions of Benoit and Krishna (1987).

Keywords: Commitment, efficiency, self-enforcing agreement, repeated games.

JEL Classification Numbers: C7

*I would like to thank Ralph Bayer, Catherine de Fontenay, Claudio Mezzetti, and Tristan Tomala for numerous and helpful conversations on this project, and Sophie Bade and Guillaume Haeringer, my co-authors of the paper "Bilateral commitment," for many insightful discussions we had on the value of commitment in games. This paper partly answers questions raised during our discussions. I am also most grateful to two anonymous referees for their comments on the paper. Most examples have been solved with GAMBIT [15].

†Department of Economics, Astley Clarke Building, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom. `l.renou@le.ac.uk`

1 Introduction

The power to commit oneself to some courses of actions or principles is undoubtedly one of the most pervasive idea in social sciences. An impressive collection of thoughtful examples illustrating the idea that it might be beneficial for individuals to constrain their behavior is found in Schelling 's seminal contribution, *The Strategy of conflict* (1960). Naturally, if individuals have the power to *jointly* commit, efficiency obtains. The intuition being that if individuals have the power to sign binding agreements, they can simply sign a contract that specify all of them to play an efficient outcome.

In a wide range of circumstances, however, players cannot sign binding agreements. There might exist no third party to enforce promises, or actions might not be verifiable. Relationship between sovereign states or international organizations is arguably the leading example of such a situation. While the International Court of Justice of the United Nations is theoretically a third party in international law, only fifty-nine states have accepted compulsory jurisdiction by July 1996. Moreover, most states accepting compulsory jurisdiction did so only with reservations and conditions. Some states that had earlier accepted compulsory jurisdiction later withdrew their acceptance when confronted with decisions, or the prospect of decisions, that they disliked.[1] In these circumstances, the power of the international court of justice to reach enforceable decisions in serious political disputes is rather limited. Yet, sovereign states have the power to *unilaterally* commit themselves to some courses of actions or principles.[2] Can efficiency obtain through unilateral commitments? What can be implemented by unilateral commitments? The purpose of this paper is to bring these questions to a close scrutiny in general strategic situations.

The approach that we follow here is to embed a game, the status quo game, in a larger game in which players can commit to some of their pure strategies in the first stage, and play the altered game in the second stage.

[1] For instance, Israel gave notice on 19 November 1985 of termination of her acceptance of such jurisdiction effective 21 November 1986; and the United States gave similar notice on 7 October 1985, effective 7 April 1986.

[2] Sovereign states can use their own laws or constitution as commitment devices.

We call the two-stage game thus defined a *commitment game*. In other words, a commitment game is a two-stage game in which the game played in the second-stage is *endogenously* determined by the commitment in the first stage. We explore how the ability to commit affects the equilibrium payoffs of the game.

Since the notion of commitment is at the heart of the paper, let us briefly dwell on it. Schelling (1960) considers several notions of commitment: "unconditional commitment," "fractional threat," or "conditional commitment." An unconditional commitment is "*a definite commitment to* ***a*** *pure strategy*," (Schelling, p184), regardless of the strategies of others, while a conditional commitment is a commitment to a pure strategy contingent on the strategies played by others. Both notions of commitment explicitly assume that players commit to a unique strategy, therefore leaves nothing open to change. In this paper, we adopt a broader view: players can commit to any subset of pure strategies.[3] Put differently, in our paper, a commitment rules out some strategies, but does not rule in. We also assume that players cannot commit to mixed strategies i.e., they cannot use "fractional threats." In our mind, commitments to mixed strategies are more difficult to communicate, and are therefore less likely to be understood and properly acknowledged as commitments. Finally, commitments are assumed to be perfect, irrevocable and perfectly observable. In the last section, we discuss how our results are likely to change if we relax some of our assumptions on the commitment technology.

Our main result is a complete characterization of the set of implementable strategies in commitment games, that is, strategies of the status quo game that are subgame perfect equilibrium outcomes. Using our characterization, we then show that pure Nash equilibria of the status quo game are always implementable. Intuitively, if all players but one commit to their equilibrium strategies, then the highest payoff the remaining player can get is his equilibrium payoff, hence committing to his equilibrium strategy is a best reply. This result already contrasts with the work of Jackson and Wilkie (2005). They consider a two-stage game, in which players can commit to contingent

[3] See chapter 8 of Schelling (1960) for a related discussion on commitments that leave something open to change.

transfers in the first stage, and play the altered game in the second stage. In particular, they show that even if the status quo game has an efficient pure Nash equilibrium, side contracting can change the set of equilibria in such a way that all equilibria of their two-stage game are then inefficient. In a commitment game, since all the (pure) equilibria of the status quo game are implementable, efficiency is preserved if the status quo game has efficient pure Nash equilibria. However, efficiency might not always result as the outcome of a commitment game. For instance, consider the prisoner's dilemma game:

	a	b
a	$3,3$	$1,4$
b	$4,1$	$2,2$

.

None of the efficient profiles (a,a), (a,b) or (b,a) are implementable. Consider the profile (a,a). If each player unilaterally commits to $\{a\}$ in the first stage of the commitment game, the induced game has $\{a\}$ as the unique strategy for each player, hence (a,a) is its unique equilibrium. However, given the commitment of player 1 to $\{a\}$, player 2 can deviate to $\{b\}$; the induced game has then strategy a for player 1 and strategy b for player 2, and its unique equilibrium (a,b) is a profitable deviation for player 2. The problem here is that the commitments needed to implement the efficient outcome are not self-enforcing. By committing to $\{a\}$, a player loses all flexibility to punish a deviating player. We will see that it is important for commitments to leave something open to change.

In order to preview more of our results, let us turn to another example. The game $G0$, below, is the status quo game, that is, the game played if players do not commit.[4]

	a	b	c	d
a	$4,4$	$0,0$	$0,3$	$-2,1$
b	$0,0$	$1/2,1/2$	$6,0$	$-1,1$
c	$0,0$	$3,2.1$	$5,3$	$1,0$
d	$5,-1$	$2,0$	$0,0$	$2,2$

,

$G0$

[4]Or, equivalently, if they commit to $\{a,b,c,d\}$, i.e., they commit not to commit.

The game $G0$ has a unique equilibrium (d, d) in pure strategies. As already argued, the profile (d, d) is implementable. Let us show that the efficient profile (a, a) is also implementable. Suppose that both players commit to $\{a, b\}$ in the first stage. The induced game is then:

	a	b
a	$4, 4$	$0, 0$
b	$0, 0$	$1/2, 1/2$

,

with two pure equilibria (a, a), (b, b), and assume that players coordinate on (a, a). Can a player profitably deviate from his commitment? The answer is no. For instance, suppose that player 1 changes his commitment in the first stage to $\{c, d\}$. The induced game has a unique equilibrium (c, b) with a payoff of 3 to player 1, hence this deviation is not profitable. The key observation is that a player does not commit to a only, but commits to a and b. Keeping b as part of his available strategy is a credible threat to punish a deviating player: If player 1 contemplates playing (d, a) by committing to $\{d\}$, player 2 punishes him by playing b. This is akin to "min-maximizing" a deviating player in the repeated game literature.

We can also show that the profiles (c, c) and (c, b) are implementable. To implement the profile (c, c), player 1 commits to $\{c\}$ and player 2 to $\{a, b, c, d\}$. Since player 2's strategy c is a best reply to c, player 2 has clearly no incentive to deviate given the commitment of player 1 to $\{c\}$. As for player 1, c maximizes his "leader" payoff, that is, the payoff to player 1 when he plays any strategy and player 2 best-replies to this strategy. That is, (c, c) is the outcome of the sequential game in which player 1 moves first, and player 2 second. This coincidence is not accidental: Theorem 1 and Lemma 1 state that to be implementable, a profile has to give each player at least the payoff he gets if he were a first mover in a certain sequential game.

Finally, we discuss at some length one special case of our model: the reduced forms of finitely repeated games. We show that the power to commit in finitely repeated games can overcome the lack of equilibrium payoffs of the constituent game being different from the "min-max" payoffs (See Benoit and Krishna (1987)). Remarkably enough, in the finitely repeated prisoner's dilemma (see above), the finite repetition of the efficient profile (a, a) is implementable.

Related literature. The literature on endogenous timing in games e.g., Hamilton and Slustky (1990, 1993), van Damme and Hurkens (1996), is closely related to our work. These authors analyze two-player commitment games in which players can either commit to a single action or not at all; commitments are unconditional.[5] They show that the equilibrium outcomes of sequential and simultaneous moves two-player games emerge as equilibrium outcomes of a two-player commitment game, hence endogenizing the order of moves. The present work differs from this literature in two important aspects. First, commitments are not restricted to single actions and the number of players is not limited to two. Second, our purpose is not to endogeneize the timing of moves in n-player finite games, but to explore how the ability to commit affects the equilibrium payoffs. To endogeneize the order of moves, a different kind of commitment games has to be considered: commitment games with at least as many periods as the number of players so that all possible orders of moves are feasible, at least in principle. Another related paper is Bade et al. (2006), who consider how the ability to commit affects equilibrium payoffs in two-player games with closed intervals of the real line as action spaces and continuous and strictly quasi-concave payoff functions. Assuming that commitments have to be closed subintervals of the original action space, they show that a profile is implementable if and only if it is implementable by a *simple* commitment. In a simple commitment, one player commits to a single action and the other commits to a subset of actions that contains his best-reply to the commitment of his opponent. Their characterization does not hold in our general context, however. Chou and Geanakoplos (1988) and García-Jurado and González-Díaz (2006) are two papers on commitment in repeated games. Chou and Geanakoplos show that if a unique player has the power to commit in the last period of a finitely repeated game, then any individually rational and feasible payoff vector is implementable. However, their result only hold for continuous games satisfying some differentiability assumptions. Moreover, commitments are exogenously

[5] A notable exception is Romano and Yildirim (2005), who consider commitments to lower bounds; all strategies above the lower bounds remaining available.

given in their model, while they are endogenously determined in our model. More closely related, García-Jurado and González-Díaz (2006) also consider commitment games based on finitely repeated games and study their virtual subgame perfect equilibria, a weakening of subgame perfection. (See Section 5 for a more extensive discussion.) Finally, Kalai et al. (2007) studies commitment games in which players can use *conditional commitments*, and show that a folk theorem obtains.

The paper is organized as follows. Section 2 describes commitment games. In Section 3, we completely characterize the equilibrium outcomes, while Section 4 uses our complete characterization as a starting point and presents further results. Section 5 considers the case of repeated games. Finally, Section 6 offers a discussion of our results and possible extensions. All proofs are in the Appendix.

2 Commitment games

A set of n players, $N := \{1, \ldots, n\}$, interact in a two-stage process. We first provide an informal description of the process.

Stage 1: Players simultaneously announce a set of *pure* strategies they commit to. That is, each player binds himself to a set of pure strategies.

Stage 2: Given the commitments in the first stage, players choose strategies in the game induced by their commitments.

A *commitment game* is thus a two-stage game, in which players commit to a subset of pure strategies in the first stage, and play the induced game in the second stage. There are many ways for players to commit to a set of strategies in the first stage, ranging from irreversible investment, reputation, legal contracts, transfer of payoffs, etc. (See Schelling (1960, Chapter 3).) In this paper, we simply assume that players have access to a perfect commitment technology without detailing what the technology is. As mentioned in the Introduction, we also assume that players cannot commit to mixed strategies. Even though commitments to mixed strategies are technically feasible (for this, it suffices to commit to randomization devices and delegate the play to a trustworthy party), we exclude this situation as, in our mind, it is difficult for a player to convince his opponents that he is committed to a randomization device. Therefore, such "randomized" commitments are likely to not be acknowledged as commitments. We refer the reader to the discussion section for more on these issues. Let us now turn to a formal description of the game. The first element of a commitment game is the status quo game, i.e., the game played if no player commits.

Status quo game. The status quo game is the strategic-form game $G := \langle N, (u_i, Y_i)_{i \in N} \rangle$ with Y_i the finite set of strategies of player $i \in N$, and $u_i : \times_{i \in N} Y_i \to \mathbb{R}$, the payoff function of player i. Let $Y := \times_{i \in N} Y_i$, $Y_{-i} := \times_{k \in N \setminus \{i\}} Y_k$, $Y_{-ij} := \times_{k \in N \setminus \{i,j\}} Y_k$. We write $\Delta(Y_i)$ for the set of mixed strategies of player i, and let $\Delta = \times_{i \in N} \Delta(Y_i)$, $\Delta_{-i} = \times_{k \in N \setminus \{i\}} \Delta(Y_k)$. Lastly, we denote by y_{-i}, y_{-ij}, y, μ_i, μ_{-i}, and μ generic elements of Y_{-i}, Y_{-ij}, Y, $\Delta(Y_i)$, Δ_{-i}, and Δ, respectively. The payoff to player i given a profile of mixed strategies $\mu = (\mu_i, \mu_{-i})$ is given by:

$$U_i(\mu_i, \mu_{-i}) = \sum_{y \in Y} \prod_{j \in N} \mu_j(y_j) u_i(y).$$

Commitments. Players can commit to subsets of their pure strategies. We denote $\mathcal{Y}_i$ the collection of all possible commitments a player can make, that is $\mathcal{Y}_i = 2^{Y_i} \setminus \emptyset$. Let $\mathcal{Y} = \times_{i \in N} \mathcal{Y}_i$. We can now formally define a commitment game.

Commitment games. The commitment game $\Gamma(G)$ (for short, Γ) is the two-stage game, in which each player commits to some $X_i \in \mathcal{Y}_i$ in the first stage, and plays $x_i \in X_i$ in the second stage. A pure strategy in Γ is a pair (X_i, σ_i) with $X_i \in \mathcal{Y}_i$ and $\sigma_i : \mathcal{Y}_i \times \mathcal{Y}_{-i} \to Y_i$ with $\sigma_i((X_i, X_{-i})) \in X_i$ for all $(X_i, X_{-i}) \in \mathcal{Y}_i \times \mathcal{Y}_{-i}$. Behavioral strategies are defined as usual. Subgame perfection is our solution concept. We say that a strategy profile $x^* \in Y$ of the status quo game is *implementable* if there exists a subgame perfect equilibrium of Γ where commitment X^* is made in the first stage, and x^* is played in the second stage (on the equilibrium path). Most of the paper is concerned with the implementation of pure strategies.

3 A complete characterization

For any commitment $X \in \mathcal{Y}$, denote $G(X)$ the game induced by the commitment X, that is, $G(X)$ is the strategic-form game $\langle N, (X_i, u_i)_{i \in N} \rangle$. Subgame perfection requires that following any commitment $X \in \mathcal{Y}$, players play a Nash equilibrium of $G(X)$.[6] Denote $NE(X_i \times X_{-i})$ the set of equilibria in pure and mixed strategies of the game $G(X_i \times X_{-i})$. As a preliminary observation, note that since a commitment game is a finite extensive-form game, it has at least one subgame perfect equilibrium. However, an equilibrium might not be in pure strategies as the following example shows. In all the examples that follow, player 1 is the row player and player 2 the column player.

Example 1. *No equilibrium in pure strategies.* Consider the game $G1$:

	a	b	c
a	$1, 0$	$6, -1$	$1, -2$
b	$3, -3$	$2, -2$	$0, -1$

$G1$

The game $G1$ has two mixed equilibria: $((\frac{1}{2}, \frac{1}{2}), (\frac{2}{3}, \frac{1}{3}, 0))$ with equilibrium payoffs of $(\frac{8}{3}, -\frac{3}{2})$, and $((\frac{1}{2}, \frac{1}{2}), (\frac{1}{3}, 0, \frac{2}{3}))$ with equilibrium payoffs of $(1, -\frac{3}{2})$. The commitment game $\Gamma(G1)$ has no subgame perfect equilibrium in pure strategies, but has 14 equilibria in behavioral strategies.[7] Moreover, some equilibria involve randomization in the first stage. For instance, there is an equilibrium, in which player 1 randomizes between $\{a\}$ and $\{a, b\}$ with probabilities $\frac{1}{3}$ and $\frac{2}{3}$ in the first stage, and player 2 plays $\{b\}$. Equilibrium payoffs are 6 and -1, respectively.

Suppose that player i can only deviate to singletons in the first stage. We can then consider the highest payoff $v_i(X_{-i})$ that player i can secure himself by committing to singletons when his opponents commit to X_{-i}, that is,

$$v_i(X_{-i}) := \max_{x_i \in Y_i} \min_{\mu \in NE(\{x_i\} \times X_{-i})} U_i(\mu). \tag{1}$$

[6] A (pure or mixed) Nash equilibrium of $G(X)$ exists for any $X \in \mathcal{Y}$ as $G(X)$ is a finite strategic-form game.

[7] Since the game $G1$ has no pure equilibria.

for any X_{-i}.[8] Note that if player i commits to the singleton $\{x_i\}$ and his opponents to X_{-i}, the equilibria of the induced game $G(\{x_i\} \times X_{-i})$ are equivalent to the equilibrium outcomes of an extensive-form game in which player i moves first and plays x_i, and the other players move simultaneously after having observed the move of player i, and choose $x_{-i} \in X_{-i}$. It is akin to a first move in leader-followers games. We call $v_i(X_{-i})$ the "leader" payoff.

Example 2. *Prisoner's dilemma.* To get some intuition for the above definition, consider the Prisoner's dilemma game $G2$.

	a	b
a	$3, 3$	$1, 4$
b	$4, 1$	$2, 2$

$G2$

It is easy to verify that $v_i(\{a\}) = 4$, $v_i(\{a, b\}) = 2$, and $v_i(\{b\}) = 2$ for all players $i \in N$. The following theorem provides a necessary condition for a profile of strategies to be implementable.

Theorem 1 *A strategy profile x^* is implementable by commitment X^* only if x^* is an equilibrium of $G(X^*)$, and $u_i(x^*) \geq v_i(X^*_{-i})$ for all player $i \in N$.*

The "proof" of Theorem 1 is as follows. First, as already mentioned, if the profile x^* is not an equilibrium of $G(X^*)$, then clearly x^* is not implementable by X^*. Second, suppose that x^* is implementable by commitment X^*, but $u_i(x^*) < v_i(X^*_{-i})$ for at least one player i, say player 1. Let us construct a profitable deviation for player 1. Suppose that player 1 commits to the strategy $\hat{x}_1$ in the first-stage of the commitment game with $\hat{x}_1$ a strategy that yields the payoff $v_1(X^*_{-1})$. The induced game is $G(\{\hat{x}_1\} \times X^*_{-1})$ and, by construction, the worst equilibrium of this induced game from player 1's perspective gives to player 1 a payoff of $v_1(X^*_{-1}) > u_1(x^*)$, a profitable deviation.

Theorem 1 highlights the link between commitment and first move in games: A strategy profile is implementable by the commitment X^* only if it

[8]Note that the minimum and the maximum are well-defined since the set of equilibria of a finite game is compact and Y_i is discrete.

gives to each player i a payoff as least as high as the payoff he would get if he were the first mover in a sequential game, where each of his opponents moves simultaneously after him and has action set X_j^*, $j \neq i$. The next section discusses in more details the connection between commitment in games and order of moves.

Before going further, observe that the definition of the "leader" payoff only considers deviations from X_i^* to singletons $\{x_i\}$. This restriction on the possible deviations is inconsequential, however, as long as any game $G(X_i \times X_{-i}^*)$ has a pure Nash equilibrium. In effect, any pure Nash equilibrium $\hat{x}$ of $G(X_i \times X_{-i}^*)$ is also a pure Nash equilibrium of the game $G(\{\hat{x}_i\} \times X_{-i}^*)$.[9] However, the condition stated in Theorem 1 is not sufficient if some games $G(X_i \times X_{-i}^*)$ do not have a pure Nash equilibrium, as the following example shows.

Example 3. *Necessary but not sufficient condition.*

	a	b	c
a	$-1/2, -1/2$	$-1, -1$	$1, -2$
b	$-1, 1$	$1, -1$	$-1, 1$
c	$1, -1$	$-1, 1$	$3, 3$

$G3$

The game $G3$ has a unique equilibrium (c, c). Consider the commitment to $\{a, b\}$ by each player. The game induced by these commitments has a unique equilibrium (a, a) with payoffs $(-1/2, -1/2)$. Moreover, we have that $v_i(\{a, b\}) = -1/2$ for each player i, hence our necessary condition is satisfied. However, (a, a) is not implementable. If player 1 deviates to $\{b, c\}$, the induced game is a game of matching pennies with equilibrium payoffs of 0 for each player, hence a profitable deviation for player 1. We can now turn to a necessary and sufficient condition for a profile of strategies to be implementable.

[9] By contradiction. Suppose that $\hat{x}$ is not a Nash equilibrium of $G(\{\hat{x}_i\} \times X_{-i}^*)$. It follows that there exists a player $j \neq i$ who has a profitable deviation $x_j \in X_j^*$ with $u_j(x_j, \hat{x}_i, \hat{x}_{-ij}) > u_j(\hat{x}_j, \hat{x}_i, \hat{x}_{-ij})$, a contradiction with $\hat{x}$ an equilibrium of $G(X_i \times X_{-i}^*)$.

Let us define the modified "leader" payoff as follows:

$$\overline{v}_i(X_{-i}) := \max_{X_i \in \mathcal{Y}_i} \min_{\mu \in NE(X_i \times X_{-i})} U_i(\mu). \tag{2}$$

The modified "leader" payoff is the highest payoff that player i can secure himself by committing to any subset of actions when his opponents have committed to X_{-i}. Note that $\overline{v}_i(X_{-i}) \geq v_i(X_{-i})$ for any $X_{-i} \in \mathcal{Y}_{-i}$.

Theorem 2 *A strategy profile x^* is implementable by commitment X^* if and only if x^* is an equilibrium of $G(X^*)$ and $u_i(x^*) \geq \overline{v}_i(X^*_{-i})$ for all player $i \in N$.*

It is easy to see that it is necessary for the modified leader payoffs $\overline{v}_i(X^*_{-i})$ to be no higher than $u_i(x^*)$ for each player i in order to implement x^* by the commitment X^*. Let us argue that this condition is also sufficient. Since x^* is assumed to be an equilibrium of $G(X^*)$, we only need to consider what happens if a player commits to a different set of strategies. If one player deviates, assume that the worst equilibrium for that player will be played in the induced (proper) subgame. From the definition of modified leader payoffs, it follows that the payoff for the deviating player i will be no more than $\overline{v}_i(X^*_{-i})$. Since $u_i(x^*) \geq \overline{v}_i(X^*_{-i})$, the deviation is not improving. If several players deviate, let them play any equilibrium of the induced game. Hence, we can construct strategies to implement any profile x^* that satisfies the above condition.

Two further remarks are worth noting. First, although Theorem 2 is stated for pure strategies, it also applies to mixed strategies.[10] Formally, a (mixed) strategy profile μ^* is implementable by commitment X^* if and only if μ^* is an equilibrium of $G(X^*)$ and $U_i(\mu^*) \geq \overline{v}_i(X^*_{-i})$ for all player $i \in N$. Second, equilibrium payoffs in any commitment game $\Gamma(G)$ are bounded from below by the "min-max" payoff $\overline{\omega}_i$ of G, that is, $u_i(x^*) \geq \overline{\omega}_i := \min_{\mu_{-i} \in \Delta_{-i}} \max_{\mu_i \in \Delta_i} U_i(\mu_i, \mu_{-i})$ for any implementable profile x^*, for all players. (See Corollary 2 in Appendix.)

It follows from Theorem 2 that to check whether a profile of strategies x^* is implementable requires two steps. The first step consists in enumerating

[10]Theorem 2 is stated for pure strategies as it is the main focus of the paper.

all games $G(X^*)$ for which x^* is a Nash equilibrium. Fix any $G(X^*)$ with x^* as one of its equilibrium. The second step consists then in enumerating *all* Nash equilibria of $G(X_i \times X^*_{-i})$ for any $X_i \in \mathcal{Y}_i$, for any player $i \in N$, and to compare the lowest Nash payoff from player i's perspective in any of those games with $u_i(x^*)$. McKelvey and McLennan (1996) survey various algorithms to enumerate all Nash equilibria of finite games; the software GAMBIT implements them. However, whether there exists a *polynomial-time* algorithm is not known. Therefore, the computational complexity of enumerating all implementable profiles of strategies of commitment games is not known.[11]

With added structure, however, the task of checking whether a profile is implementable can be simplified. For instance, consider a discrete version of a duopoly game with differentiated goods. The payoff to player i if he produces $x_i \in Y_i$ and player j produces $x_j \in Y_j$ is $(1 - x_i + \gamma x_j)x_i$, with $Y_i = \{0, 1/4, 1/3, 1/2, 2/3, 3/4, 1\} = Y_j$. Goods are substitutes if $\gamma < 0$ and complements, otherwise. Note that if player i commits to X_i, player i's (restricted) best replies to any $x_j \in Y_j$ are the points in X_i the closest to $(1 + \gamma x_j)/2$.[12] For $\gamma = -1$, the set of implementable profiles of strategies is $\{(1/3, 1/3), (1/2, 1/4), (1/4, 1/2)\}$. For instance, the profile $(1/2, 1/4)$ is implementable by the commitment $(\{1/2\}, Y_j)$. Since $1/4$ is the best-reply of player j to $1/2$, we have that $\overline{v}_j(\{1/2\}) = u_j(1/4, 1/2)$. Moreover, we have that $\overline{v}_i(Y_j) = \max_{x_i \in Y_i} u_i(x_i, BR_j(x_i)) = u_i(1/2, 1/4)$ where BR_j is the "worst" selection of player j's best-reply map $\overline{BR}_j : Y_i \to \mathcal{Y}_j$ from player i's perspective.[13] Let us now show that the efficient profile $(1/4, 1/4)$ is not implementable. First of all, to implement $(1/4, 1/4)$, it has to be a Nash equilibrium of some game $G(X^*)$. Since the point $1/3$ is closer to $(1 - 1/4)/2 = 3/8$, i's best-reply to $1/4$, than $1/4$, we should have $X^*_i \cap \{1/3\} = \emptyset$ for each player i. For otherwise, $(1/4, 1/4)$ cannot be a Nash equilibrium of $G(X^*_i \times X^*_j)$. Second, consider any commitment (X^*_i, X^*_j) such that $1/4 \in$

[11]Equivalently, we have to find for any player i, for any $X_i \in \mathcal{Y}_i$, a Nash equilibrium of $G(X_i \times X^*_{-i})$ with a payoff smaller than $u_i(x^*)$ to player i. This problem is $\mathcal{NP}$-hard (Gilboa and Zemel (1989)).

[12]Since u_i is strictly quasi-concave in x_i and symmetric.

[13]Precisely, for any x_i, we select an element $BR_j(x_i)$ of $\overline{BR}_j(x_i)$ such that $u_i(x_i, BR_i(x_i)) \leq u_i(x_i, x_j)$ for any $x_j \in \overline{BR}_j(x_i)$.

$X_i^* \subset Y_i \backslash \{1/3\}$ and $1/4 \in X_j^* \subset Y_j \backslash \{1/3\}$. We have that $(1/4, 1/4)$ is a Nash equilibrium of $G(X_i^* \times X_j^*)$. Suppose that player i changes his commitment to $\{1/3\}$: the induced game is $G(\{1/3\} \times X_j^*)$. This game has a unique Nash equilibrium $(1/3, 1/4)$ since $1/4$ is the closest point to $(1 - 1/3)/2 = 1/3$ in the set $Y_j \setminus \{1/3\}$. The associated payoff to player i is $5/36$, henceforth $\overline{v}_i(X_j^*) \geq 5/36$. Since player i's payoff to $(1/4, 1/4)$ is $1/8 < 5/36$, it follows from Theorem 2 that $(1/4, 1/4)$ is not implementable. A similar reasoning applies to the other profiles. We now consider the case $\gamma = 1/2$ and show that the efficient profile $(1, 1)$ is implementable with a payoff of $1/2$ to each player. As above, since the points $2/3$ and $3/4$ are closer to $(1+0.5)/2 = 3/4$ than 1 is, we should have that $X_i^* \subset Y_i \setminus \{2/3, 3/4\}$ for $(1, 1)$ to be an equilibrium of $G(X^*)$. Also, observe that only two profiles $(3/4, 1)$ and $(2/3, 1)$ are strictly preferred to $(1, 1)$ by player 1. Thus, if we can guarantee that these profiles are not equilibria of any game $G(X_1 \times X_2^*)$, we can insure that $\overline{v}_i(X_2^*) \leq 1/2$. To this end, note that $1/2$ is closer to $(1 + 0.5(3/4))/2 = 11/16$ and $(1 + 0.5(2/3))/2 = 2/3$ than 1 is. Therefore, if player 2 includes $1/2$ in his commitment X_2^*, we can be sure that $(3/4, 1)$ and $(2/3, 1)$ won't be Nash equilibria of any game $G(X_1 \times X_2^*)$. By symmetry, the same argument apply to player 2. It follows then from Theorem 2 that the efficient profile $(1, 1)$ is implementable, by the commitment to $\{1/2, 1\} \times \{1/2, 1\}$ for instance.

As another example, if G is the reduced form of a finitely repeated game, we can obtain sufficient conditions for a folk theorem to obtain (see section 5).

4 Further results

In this section, we take Theorem 2 as a starting point to discuss two important issues about commitment in games: efficiency and order of moves. We first discuss the connection between commitment and order of moves in games.

4.1 Commitment and order of moves

Our first result on the order of moves, Corollary 1, states that pure equilibrium outcomes of the status quo game G are equilibrium outcomes of the commitment game $\Gamma(G)$. In other words, pure equilibrium outcomes of the extensive-form game in which all players move simultaneously are implementable.

Corollary 1 *Pure equilibria of the status quo game G are implementable.*

To get some intuition on Corollary 1, suppose that for all $i \in N$, player i commits to a single strategy $\{x_i^*\}$.[14] Player i's modified "leader" payoff $\overline{v}_i(\times_{j\neq i}\{x_j^*\})$ is then $\max_{x_i \in Y_i} u_i(x_i, x_{-i}^*)$. Since, from Theorem 2, a profile x^* is implementable if and only if $u_i(x^*) \geq \overline{v}_i(\times_{j\neq i}\{x_j^*\})$ for each player $i \in N$, it immediately follows that the pure Nash equilibria of the status quo game are implementable. However, Corollary 1 does not hold for the mixed equilibria of G: mixed equilibria might or might not be implementable.[15] For an example in which mixed equilibria are implementable, consider the game of matching pennies. For an example in which mixed equilibria are not implementable, consider the game $G4$ below.

Example 4. *Mixed equilibria are not implementable.*

	a	b	c
a	$3,3$	$6,1$	$2,2$
b	$4,5$	$1,6$	$3,3$

$G4$

Game $G4$ has a unique Nash equilibrium in mixed strategies $((\frac{1}{3},\frac{2}{3}),(\frac{5}{6},\frac{1}{6},0))$ with equilibrium payoffs $\frac{7}{2}$ and $\frac{13}{3}$, respectively. However, it is not implementable since $\overline{v}_2(\{a,b\}) = 5 > 13/3$. An equilibrium of $G4$ is as follows: In the first stage, player 1 commits to $\{a,b\}$ and player 2 to $\{a,c\}$, and in the second-stage (b,a) is played (on the equilibrium path). Note that (b,a) is also the outcome of the sequential game in which player 2 moves first and

[14] This result first appears in Bade et al. (2006).

[15] If players can commit to mixed strategies, Lemma 1 holds for the mixed equilibria of G.

player 1 moves second, after having observed the move of player 2. This result is not accidental as the next lemma will show.

Consider the following assumption:

A0: For any pair of players $(l, m) \in N \times N$, any profile of strategies $y_{-lm} \in Y_{-lm}$, the two-player game $\langle \{l, m\}, (X_i, \overline{u}_i)_{i \in \{l,m\}} \rangle$ with $\overline{u}_i(x_l, x_m) := u_i(x_l, x_m, y_{-lm})$ has a pure Nash equilibrium, for any $X_i \in \mathcal{Y}_i$, $i \in \{l, m\}$.

Lemma 1 *Any equilibrium outcome of the extensive-form game in which $n-1$ players move simultaneously in a first stage, and the remaining player plays in a second stage, is implementable in the commitment game if A0 holds.*

Lemma 1 is a generalization to n-player games of Hamilton and Slustky (1990, 93) result for two-player games, which states that Stackelberg outcomes are equilibrium outcomes of commitment games.[16] Note that condition A0 cannot be dispensed with. Once again, the game of matching pennies constitutes a counter-example. It is also worth pointing out that A0 is satisfied if the status quo game G is a supermodular game (see Milgrom and Roberts (1990) or Topkis (1998)). To illustrate Lemma 1 further, let us consider the following example.

Example 5. In the game $G5$, below, player 1 chooses a row, player 2 a column and player 3 a matrix. This game has a unique equilibrium (a, a, a). Let us show that the efficient profile (a, b, b) is implementable.

	a	b
a	$1, 1, 1$	$0, 0, 0$
b	$0, 0, 1$	$3, 3, 0$

a

	a	b
a	$0, 3, 0$	$2, 2, 3$
b	$0, 0, 0$	$0, 0, 1$

b

$G5$

We first easily check that the profile (a, b, b) (with $BR_3(a, b) = b$) is the equilibrium outcome of the sequential game in which, players 1 and 2 move simultaneously in the first stage and player 3 moves in the second

[16]Note, however, that Hamilton and Slustky consider a special class of two-player commitment games, in which each player can either commit to a singleton or not at all.

stage. Second, let us show that the profile (a, b, b) can be implemented by the commitment on the part of players 1 and 2 to a and b, respectively, while player 3 does not commit. Since b is player 3's best-reply to (a, b), player 3 has clearly no incentive to deviate given the commitment of players 1 and 2. Given the commitment of player 2 to b and player 3 to a and b, if player 1 has a profitable deviation from a to b, his payoff is $u_1(b, b, BR_3(b, b)) > u_1(a, b, BR_3(a, b))$, a contradiction with (a, b, b) being an equilibrium of the sequential game. Similarly, for player 2. We also easily verify that $A0$ holds.

Let us summarize our results on commitment and order of moves. We have seen that if all players commit to a single strategy, the only implementable profiles of strategies are the pure Nash equilibria of the status quo game (Corollary 1). Moreover, Lemma 1 characterizes the profile of strategies that are implementable if $n - 1$ players commit to a single strategy and the remaining player does not commit. We might then wonder what can be implemented if $n - 2$ players commit to a single strategy and the remaining two players do not commit (or commit to subsets of strategies larger than singletons), and proceed by induction. The key observation is that those implementable profiles are equivalent to equilibrium outcomes of extensive-from games in which $n - 2$ players move simultaneously in a first stage, and the remaining players also move simultaneously in a second stage. Unfortunately, for those profiles to be implementable, far stronger assumptions than A0 have to be imposed. We therefore do not pursue this issue, here.

4.2 Commitment and efficiency

Another interesting issue is whether the power to commit is conducive to efficiency. We have already seen that the power to commit does not always help to implement efficient outcomes (see $G2$). In this section, we discuss the trade-off between commitment and flexibility to punish opponents that underlies the implementation of efficient outcomes. Consider the game G_δ below with $\delta < 1$.

	a	b	c
a	$3,3$	$1,2$	$0,4$
b	$2,1$	$0,0$	$\delta,2$
c	$4,0$	$2,\delta$	$1,1$

G_δ

Is the efficient profile (a, a) implementable? First, observe that (a, a) is an equilibrium of the game $G(\{a, b\} \times \{a, b\})$, and dominates all other equilibria of $G(\{a, b\} \times \{a, b\})$. Second, we have that $\overline{v}_i(\{a, b\}) = 4$ if $\delta < 0$, and $\overline{v}_i(\{a, b\}) = 3$ if $\delta \geq 0$ for both players. The efficient profile (a, a) is thus implementable by commitment $\{a, b\} \times \{a, b\}$ if $\delta \geq 0$, but is not, otherwise. The intuition behind this result is that if $\delta \geq 0$, by committing to $\{a, b\}$, a player retains enough flexibility to threat (off-equilibrium) deviations. In other words, if player i deviates to any commitment which includes c, player $j \neq i$ can credibly punish the deviating player by playing b, which leads to a payoff of 2 instead of 3 for the deviating player. Thus, if we want to implement an efficient profile of strategies x^*, there is a trade-off between enough flexibility to punish off-equilibrium deviations and enough inflexibility to implement x^* as an equilibrium of the game induced by the commitment. This is the main principle behind the implementation of efficient outcomes. Note that if $\delta < 0$, none of the efficient profiles (a, a), (a, c) and (c, a) are implementable. Lastly, it is worth pointing out that G_δ is dominance solvable regardless of $\delta < 1$, hence whether efficient profiles are implementable or not is a *cardinal* property of the game under consideration. For the class of finitely repeated games, Section 5 presents sufficient conditions for efficient profiles of strategies to be implementable.

So far, we have seen through many examples that the possibility to commit in games often enlarges the set of equilibrium payoffs (with respect to the status quo) to the extent that efficient profiles might even be implementable. To conclude this section, we aim at answering the following question:[17] when the possibility to commit does not enlarge the set of equilibrium payoffs? While a complete answer awaits future research, we can isolate a class of games for which commitment does not help. Consider the class of status quo

[17] I thank Eyal Winter for having suggested to me this question.

games G for which the strategy set Y_i is totally ordered by the strict dominance relation for each player $i \in N$.[18] We can then show that the set of implementable strategies is the set of Nash equilibria. Commitment does not help for this class of games.[19] Finally, note that the assumption of strategy sets being totally ordered by the strict dominance relation is a very strong requirement. However, it cannot be easily relaxed. For instance, consider the game G_δ with $\delta = 0$, and observe that c strictly dominates a and b, and a weakly dominates b. The efficient profile (a, a) is implementable while it is not a Nash equilibrium of the status quo game.

5 Commitment and repeated games

In this section, we restrict our attention to a special class of status quo games: the reduced forms of *finitely* repeated games. For such games, a strategy is a complete (contingent) plan of actions, and commitments to strategies allow players to commit to a particular action after certain histories, to another after other histories, etc. In turn, this possibility of "conditioning" enables the players to implement a large set of action profiles.

Formally, we first define the constituent game g, that is, the game repeatedly played. Let g be the strategic-form game $\langle N, (w_i, A_i)_{i \in N} \rangle$ with A_i the finite set of actions of player $i \in N$, and $w_i : A := \times_{i \in N} A_i \to \mathbb{R}$, the payoff function of player i. We assume that g is generic, that is, for any $a \in A$, $a' \in A \setminus \{a\}$, $w_i(a) \neq w_i(a')$. We then define G^T as the T-repetition of the game g with perfect monitoring. A (non-initial) history at stage $t+1$ is an

[18] We say that the strategy set Y_i is totally ordered by the strict dominance relation if for any two different strategies x_i and x_i' in Y_i, either x_i strictly dominates x_i', or x_i' strictly dominates x_i. For instance, the set of strategies in the Prisoner's dilemma game is totally ordered by the strict dominance relation.

[19] Note that Romano and Yildirim (2005) show that *only* the Cournot-Nash outcomes are implementable if the Stackelberg-leader outcomes are "smaller" than the Cournot-Nash outcomes (their Proposition 2). Their proposition does not hold in our more general framework. Indeed, for the class of two-player supermodular games –the class considered in Romano and Yildirim– the Stackelberg outcomes are always implementable by the commitment of one player (the leader) to a singleton and no commitment for the other player (the follower). See Lemma 1.

element of $\times_{s=0}^{t} A^s$, that is, a sequence of t profiles of actions. We define the initial history A^0 to be $\{\emptyset\}$. Let $H := \bigcup_{t=0}^{T} A^t$ be the set of histories. A (pure) strategy s_i for player i is a map from the set of non-terminal histories to the set of actions, that is, $s_i : \bigcup_{t=0}^{T-1} A^t \to (A_i)^T$. Denote S_i the set of all pure strategies for player i in the game G^T and $S = \times_{i \in N} S_i$. A strategy profile $s \in S$ recursively and uniquely determines a sequence of actions $(\emptyset, s(\emptyset), s((\emptyset, s(\emptyset))), \dots)$, that is, the path generated by the strategy profile s. Let $(a^1, \dots, a^T)$ the terminal history induced by the profile of strategies (s_i, s_{-i}), player i's payoff in G^T is

$$u_i(s_i, s_{-i}) = \frac{1}{T} \sum_{t=1}^{t=T} w_i(a^t).$$

The status quo game is thus the strategic-form game $\langle N, (S_i, u_i)_{i \in N} \rangle$, that is, the reduced form of the finitely repeated game G^T. With a slight abuse of notations, we also denote the status quo game by G^T. The commitment game $\Gamma(G^T)$ is defined as in Section 2. Given a commitment $R_i \subseteq S_i$, we can define $X_i(h^t)$ the set of actions available to player i at (non-terminal) history h^t as

$$X_i(h^t) := \{a_i \in A_i : a_i = s_i(h^t) \text{ for some } s_i \in R_i\}.$$

Let $X(h^t) = \times_{i \in N} X_i(h^t)$ and $X_{-i}(h^t) = \times_{j \in N \setminus \{i\}} X_j(h^t)$. Lastly, we denote D the set of histories in the commitment game (D should not be confused with H, the set of histories of G^T). Histories are defined as follows: $d^0 = \{\emptyset\}$ is the initial history, $d^1 = (d^0, R)$ with $R \subseteq S$, $d^2 = (d^1, x^1)$ with $x^1 \in X(h^0)$, $d^3 = (d^2, x^2)$ with $x^2 \in X(h^1)$, $h^1 = (h^0, x^1)$, etc.[20] Observe that any commitment R in the first stage induces a strategic-form game $G^T(R)$, which is not necessarily the reduced form of a repeated game. Yet, if each player commits to the same subset of actions $B_i \subseteq A_i$ after any history h, the induced game is the reduced-form of a repeated game.

5.1 A Folk theorem

A natural conjecture is that any feasible and individually rational payoff is equilibrium payoff (or, at least, can be arbitrary approximated) of the

[20] See Osborne and Rubinstein (1994) for more on extensive-form games.

commitment game $\Gamma(G^T)$. The next theorem gives such a folk theorem. Denote ω_i the "min-max" payoff to player i in pure actions.[21]

Theorem 3 *Suppose that for each player $i \in N$, there exists a profile of pure actions ${}^ie^*$ implementable by ${}^iX^*$ in $\Gamma(g)$ such that $u_i({}^ie^*) > \omega_i$. Let $U \in \mathbb{R}^n$ be any individually rational and feasible payoff vector of g. For all $\varepsilon > 0$, there exist a T^* and an equilibrium of the commitment game $\Gamma(G^T)$ whose payoff vector u satisfies $||U - u|| \leq \varepsilon$ for any $T \geq T^*$.*

Theorem 3 is reminiscent of Theorem 1 in Benoit and Krishna (1987). From Corollary 1, pure Nash equilibria of g, the constituent game, are implementable. Hence, if for each player i, the game g has a pure Nash equilibrium, which gives a payoff strictly higher than the "min-max" payoff to player i, the result follows from Theorem 1 of Benoit and Krishna (1987) and Corollary 1 (applied on G^T, this time). More generally, González-Diáz (2005) has shown that the Nash folk theorem for G^T obtains if and only if the game g is decomposable as a complete minimax-bettering ladder. Therefore, if g is decomposable as a complete minimax-bettering ladder, Theorem 3 follows from Theorem 1 (p. 106) of González-Diáz (2005) and Corollary 1. Note that this observation is essentially Proposition 2 of García-Jurado and González-Díaz (2006). However, Theorem 3 is more general. For instance, consider the game g below.[22]

	a	b
a	$2, 1$	$0, -1/2$
b	$3, -1$	$1, 0$

g

The profile (b, b) is the unique Nash equilibrium of this game. Moreover, the equilibrium payoff is identical to the "min-max" payoff. Hence, the unique Nash equilibrium of G, the finite repetition of g, is to play b in every period.[23]

[21] Formally, $\omega_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_i, a_{-i})$. A payoff is individually rational if it is above ω_i.

[22] For another example, consider Example 2 in García-Jurado and González-Díaz (2006). The profile (U, L) is implementable by the commitment $\{U\} \times \{L, M\}$ and gives a payoff of 10 to each player, while the "min-max" payoff to players 1 and 2 is 1 and 2, respectively.

[23] The game g is *not* decomposable as a complete minimax-bettering ladder.

However, if players have the power to commit, they can do better. Observe that the profile (a, a) of g is implementable by the commitment of player 1 to $\{a\}$ and player 2 to $\{a, b\}$, and gives a payoff strictly higher than the "min-max." (See Lemma 1.) The basic idea is then very simple: If there exist (self-enforcing) commitments that implement any *finite* sequence of (a, a), we can then replicate the arguments in Benoit and Krishna (1987) to approximate any individually rational and feasible payoffs' profile as an equilibrium of our commitment game. Indeed, by committing to play the implementable profile (a, a) enough times toward the end of the game, every player is deterred from deviating, if threatened with his "min-max" payoff for the remaining periods of the game after a deviation. Now, if we wish to approximate the individually rational payoff $(3/2, 1/2)$, say, it is enough, as in Benoit and Krishna, to play (a, a) and (b, b) enough times at the beginning of the game for the same numbers of periods (a, a) is played toward the end of the game. It therefore remains to show that there are (self-enforcing) commitments that implements any finite sequence of (a, a). This step is done in Appendix.

While Theorem 3 improves upon the folk theorem of Benoit and Krishna (1987) and González-Diáz (2005), the finitely repeated prisoners' dilemma does not satisfy the premise of Theorem 3 (see $G2$). In the sequel, we show that if an action profile (of g) satisfies the so-called *no-reward condition*, then it can be implemented. Before giving the formal definition, let us illustrate the no-reward condition with the prisoners' dilemma game $G2$ repeated twice. Can commitment sustain the efficient outcome $((a, a), (a, a))$? From Theorem 2, we first need to find a game $G^2(R^*)$ such that $((a, a), (a, a))$ is an equilibrium outcome of this game. Suppose that both players commit to play a in the second period if (a, a) is played in the first period and to play b if another profile is played in the first period.[24] If both players play a in the first period, their (normalized) payoff is 3 $(= (3 + 3)/2)$. If a player deviates to b in the first period, his payoff is 3. Therefore, $((a, a), (a, a))$ is an equilibrium outcome of $G^2(R^*)$. We second need to check whether $3 \geq v_i(R^*_{-i})$. Given the commitment of player 2 to R^*_2, the only outcome that gives player 1 a payoff higher than 3 is $((a, a), (b, a))$. Can player 1 induces this sequence

[24]That is, $R^*_i = S_i \setminus \{s_i \in S_i : s_i(a, a) = b, s_i(b, a) = a, s_i(a, b) = a, s_i(b, b) = a\}$.

as the equilibrium of a game $G(R_1 \times R_2^*)$. In order to do so, player 1 needs to *reward* player 2 for playing a in the first period. However, if player 2 plays b in the first period, the lowest payoff player 1 can force player 2 to in the second period is 2. Consequently, if player 2 plays b, the lowest payoff he can obtain is 3, while he would get 2 if he plays a. It is therefore impossible for player 1 to reward player 2 for playing a: (a, a) satisfies the *no-reward condition*, which we now formally define. Let $m^i \in A$ be the action profile that "min-max" player i, that is, $u_i(m^i) = \omega_i$.

Definition 1 *Let $a^* \in A$ and define τ^* as the smallest integer larger or equal to*

$$\max_{i \in N} \left(\frac{\max_{a \in A} u_i(a) - u_i(a^*)}{u_i(a^*) - \omega_i} \right).$$

*The action profiles a^* satisfies the **no-reward condition** if for all players $j \in N$ with $\{a'_j : u_j(a'_j, a^*_{-j}) > u_j(a^*)\} \neq \emptyset$, there exists a player $i \in N \setminus \{j\}$ and an integer $\tau^{**} \geq \tau^*$ such that*

$$\sum_{s=1}^{\tau^{**}} u_i(\hat{a}_j(s), a^*_{-j}) \leq \tau^{**} \min_{a_j \in A_j} u_i(a_j, m^j_{-j}), \tag{3}$$

with $(\hat{a}_j(s))_{s=1}^{\tau^{**}}$ solution of

$$\max_{\{a_j(s)\} \in A_j^{\tau^{**}}} \sum_{s=1}^{\tau^{**}} u_i(a_j(s), a^*_{-j})$$

subject to

$$\sum_{s=1}^{\tau^{**}} u_j(a_j(s), a^*_{-j}) > \tau^{**} u_j(a^*).$$

The inequality in Equation (3) has to be strict if $\{a'_i : u_i(a'_i, a^*_{-i}) > u_i(a^*)\} = \emptyset$. We say that the payoff vector $u \in \mathbb{R}^n$ satisfies the no-reward condition if there exists an action profile a^* of g that satisfies the no-reward condition, and $u = u(a^*)$.

In words, a profile of actions a^* satisfies the no-reward condition if no player j can profitably deviate from the sequence of a^* played $\tau^{**} + 1$ times when *all* the other players are committed to play a^*_{-j} for the last τ^{**} periods if a^* is played in the first period, and to "min-max" player j for the last τ^{**} periods if a^* is **not** played in the first period. The no-reward condition ensures that if player j contemplates (profitably) deviating from a^* in at least one of the last τ^{**} periods, there exists a player $i \neq j$ who will deviate from a^* in the first period. Moreover, since $\tau^{**} \geq \tau^*$, this ensures that player j does not benefit from the deviation of player i in the first period if a^* is individually rational. Returning to the game $G2$, (a, a) satisfies the no-reward condition with $\tau^* = 1$, and $\tau^{**} = 1$.[25] Indeed, the highest payoff that a player, say player 1, can give to player 2 when player 2 is committed to a is to play a for $\tau^{**} - 1$ periods and b for one period (subject to the constraint that it gives to player 1 a payoff higher than 3). This gives a payoff of $((\tau^{**} - 1)3 + 1)/\tau^{**}$ to player 2. The lowest payoff player 1 can give to player 2 when player 2 is committed to play b (the action that "min-max" player 1) is 2. Hence, if there exists a $\tau^{**} \geq 1$ such that $(\tau^{**} - 1)3 + 2 \leq \tau^{**}2$, then (a, a) satisfies the no-reward condition; $\tau^{**} = 1$ satisfies the inequality.

We say that the game g has the *common threat* property if $m^i = m^j$ for all players i and j. We write m^* for the common threat.

Theorem 4 *Suppose that the game g has the common threat property and is either a two-player game or its T-repetition, G^T, is a supermodular game. Let u be an individually rational and feasible payoff vector, which satisfies the no-reward condition. There exists a T^* and an equilibrium of the commitment game $\Gamma(G^T)$ whose payoff vector is u for any $T \geq T^*$.*

To get intuitions on Theorem 4, suppose we wish to implement the payoff from the profile a^*, and $T = \tau^{**} + 1$. Consider *trigger commitments* in which players commit to a^* for the last τ^{**} periods if a^* is played in the first period and to m^*, the common threat, for the last τ^{**} periods if a^* is not played in the first period. If all players use the trigger commitments just described, $(a^*, \ldots, a^*)$ is an equilibrium outcome of the game induced by these commitments. Indeed, if one player deviates from a^* in the first period, he receives his "min-max" payoff for the remaining τ^{**} periods. Since

[25] In fact, we can show that for any prisoners' dilemma game, the "cooperative outcome" satisfies the no-reward condition. An alternative proof of the implementation of the cooperative outcomes games was independently found in Faína et al. (1998).

$\tau^{**} \geq \tau^*$ (see Definition 1), this deviation is not profitable. If player j deviates from his trigger commitment, he can benefit from this deviation only if a^* is played in the first period following his deviation. However, since a^* satisfies the no-reward condition, there exists a player i who cannot be rewarded to play a_i^* in the first period. Hence, following any deviation by player j from his trigger commitment, there is no pure equilibrium of the induced game for which a^* is played in the first period. The common threat property then ensures that any profile $(a, m^*, \ldots, m^*)$ with $a \neq a^*$ gives a payoff to player j lower than $u_j(a^*)$, j's payoff under a^*. Finally, if G^T is a supermodular game, each induced game has a pure Nash equilibrium, and from the above arguments, a profile $(a, m^*, \ldots, m^*)$ with $a \neq a^*$ has to be played on the equilibrium path.[26] And if g is a two-player game, it cannot be part of an equilibrium to play a^* with positive probability. If $T \geq \tau^{**} + 1$, we have to slightly modify the trigger commitments to get the desired result: Players commit to a^* for the last τ^{**} periods if a^* has always been played in the first $T - \tau^{**}$ periods and to m^* for the remaining periods as soon as a deviation from a^* is observed. In other words, for any profile a^* that satisfies the no-reward condition, we can construct a trigger commitment R_i^* for each player i such that $u_i(a^*) = \overline{v}_i(R_{-i}^*)$, hence by Theorem 2, the finite repetition of a^* is implementable.

6 Discussion

This paper characterizes the strategy profiles of strategic-form games that are implementable when players can unilaterally commit to subsets of strategies, but cannot sign (jointly) binding agreements. We notably show that if one wants to implement a strategy profile x^* by the commitment X^*, the commitment X^* has to be small enough to make x^* a Nash equilibrium of $G(X^*)$, but has also to be large enough to punish eventual deviations. Through examples, we have shown that efficient profiles can be implemented. Let us discuss some of the restrictions on the commitment technology we have considered, and how our results are likely to change with altered assumptions.

[26]See Echenique (2004) for the concept of supermodular extensive-form games.

Commitment to mixed strategies. We have assumed that players cannot commit to mixed strategies. Nonetheless, it is straightforward to accommodate our analysis to account for commitments to mixed strategies. For instance, the modified version of Corollary 1 would be that all equilibria, whether pure or mixed, of the status quo game are implementable. However, some technical issues have to be considered as the commitment to a set of mixed strategies $X_i \subset \Delta_i$ might not be a compact and/or convex set, hence the induced game might have no equilibrium. If we assume that players have to commit to non-empty compact sets (of mixed strategies), the characterization of implementable profile remains literally the same.[27]

Veto power. An implicit assumption in a commitment game is that a player cannot veto the commitment of other players; commitments are unilateral. However, in many instances such as international organizations, there is a mix of unilateral commitments and the possibility of vetoes (e.g., the five permanent members of the UN security council have the power to veto any proposal on substantive matters). How does the possibility of veto affect the set of implementable profiles? To answer this question, suppose that after the commitments have been announced, players have the opportunity to veto the proposed commitments. If a commitment is vetoed, the status quo game is played. An immediate consequence is that in a commitment game with veto power, any implementable profile has to dominate a Nash equilibrium of the status quo game i.e., only Pareto-improvements are implementable. Moreover, any profile that is implementable in a commitment game without veto power and that dominates a Nash equilibrium of the status quo game is also implementable in a game of commitment with veto power. However, other profiles can be implemented. For instance, in the game G_δ with $\delta < 0$, the profile (a, a) is not implementable in a commitment game without veto power while it is implementable in a commitment game with veto power. Simply, the deviation to $\{c\}$ by either players is vetoed. Thus, the efficient profile (a, a) is implementable in a commitment with veto power (similarly, in the Prisoners' dilemma, the efficient profile is implementable). It would then be tempting to say that the mere possibility of commitment and veto

[27]By Glicksberg's theorem, any induced game has a Nash equilibrium in mixtures of mixed strategies, and the analysis of Section 3 carries on.

power is enough to implement the efficient profiles. The following shows that this intuition does not hold.

Example 6. *A commitment with veto power.*

	a	b	c	d
a	$3,3$	$0,0$	$1,4$	$0,0$
b	$0,0$	$-1,-1$	$0,0$	$0,0$
c	$4,1$	$0,0$	$2,2$	$0,3$
d	$0,0$	$0,0$	$3,0$	$1,1$

$G6$

The game $G6$ has a unique Nash equilibrium (d, d). However, none of the efficient profiles (a, a), (a, c), and (c, a) are implementable. For instance, if players 1 and 2 commit to $\{a, b\}$ in the first stage, then player 2 has an incentive to deviate to $\{c\}$ since his payoff would then be 4. Moreover, the proposal $(\{a, b\}, \{c\})$ is not vetoed as it improves upon the status quo payoffs.

Commitment to transfers. There is a fundamental difference between commitment to transfer functions as in Jackson and Wilkie (2005) and to strategies. On the one hand, if player i commits to transfer functions $t_{ij} : Y \to \mathbb{R}$, with $t_{ij}(x)$ the transfer from i to j conditional on the strategy x being played, then player j can undo the transfer by committing to $t_{ji}(x) = -t_{ij}(x)$. On the other hand, in a commitment game, the commitment to a set of strategies cannot be undone; a player cannot give back the freedom of choice to another player. The following example shows that this difference is not without consequences.

Example 7. *Commitment to transfers or strategies.*

	a	b	c
a	$3,3$	$0,0$	$1/2,4$
b	$0,0$	$-1,-1$	$2/3,2$
c	$4,1/2$	$2,2/3$	$1,1$

$G7$

Following the analysis in Jackson and Wilkie, it is easy to see that the solo payoff u_i^s is 7/2 to each player i in game $G7$. To see this, observe that if player 1 transfers 1/2 to player 2 conditional on (c, a) being played, then he can support (c, a) as an equilibrium with a payoff of 7/2 for himself and 1 for player 2. Henceforth, from Theorem 2 of Jackson and Wilkie, the profile (a, a) cannot be sustained as an equilibrium with transfers. However, if both players commit to $\{a, b\}$, then (a, a) is supportable as an equilibrium of a commitment game.

7 Appendix

Corollary 2 *Any implementable x^* is such that $u_i(x^*) \geq \overline{\omega}_i$ for all $i \in N$.*

Proof of Corollary 2 Let (X^*, μ^*) be an equilibrium outcome of $\Gamma(G)$. From Theorem 2, we have that $U_i(\mu^*) \geq \overline{v}_i(X_{-i}^*)$ for all players $i \in N$. In particular, we have

$$u_i(\mu^*) \geq \min_{\mu \in NE(Y_i \times X_{-i}^*)} U_i(\mu),$$

since Y_i is a possible deviation for player i. Let $\hat{\mu}$ be a Nash equilibrium of $G(Y_i \times X_{-i}^*)$ such that $U_i(\hat{\mu}) = \min_{\mu \in NE(Y_i \times X_{-i}^*)} U_i(\mu)$. We have

$$\begin{aligned} U_i(\hat{\mu}) &\geq \min_{\mu_{-i} \in \Delta(X_{-i}^*)} \max_{\mu_i \in \Delta_i} U_i(\mu_i, \mu_{-i}) \\ &\geq \min_{\mu_{-i} \in \Delta_{-i}} \max_{\mu_i \in \Delta_i} U_i(\mu_i, \mu_{-i}) = \overline{\omega}_i, \end{aligned}$$

since $\Delta(X_{-i}^*) \subseteq \Delta_{-i}$, which completes the proof. □

Proof of Lemma 1 For any $k \in N$, let $BR_k : Y_{-k} \to Y_k$ be a selection of player k's best-reply correspondence of the status quo game G. Define the game $g_k := \langle N \setminus \{k\}, (w_i, Y_i)_{i \in N \setminus \{k\}} \rangle$ with $w_i(y) = u_i(y, BR_k(y))$ for all $y \in \times_{i \neq k} Y_i$. In words, g_k is a game obtained from G, in which player k is "inactive" and $x_k = BR_k(x_{-k})$ for all $x_{-k} \in Y_{-k}$.

Let $k = n$ and $x_{-n}^* = (x_1^*, \ldots, x_{n-1}^*)$ a pure Nash equilibrium of g_n. Construct equilibrium strategies as follows. In the first stage, each player $i \in N \setminus \{n\}$ commits to $\{x_i^*\}$ and player n commits to Y_n. In the second stage, strategies prescribe the play of a Nash equilibrium in any induced game. In particular, player n plays $BR_n(x_1^*, \ldots, x_{n-1}^*)$ following the commitment to $\{x_{-n}^*\} \times Y_n$.

Since player n best replies to x_{-n}^*, he has no incentive to deviate from his commitment given the commitment of his opponents to x_{-n}^*. Let us now consider player $1 \in N \setminus \{n\}$. Since $(x_1^*, \ldots, x_{n-1}^*)$ is an equilibrium of g_n, we have that

$$u_1(x_1^*, x_2^*, \ldots, x_{n-1}^*, BR_n(x_1^*, \ldots, x_{n-1}^*)) \geq u_1(x_1, x_2^* \ldots, x_{n-1}^*, BR_n(x_1, x_2^* \ldots, x_{n-1}^*)),$$

for all $x_1 \in Y_1$. By contradiction, suppose that player 1 has a profitable deviation to X_1. The induced game is $G(X_1 \times \{x_2^*\} \times \ldots \{x_{n-1}^*\} \times Y_n)$. For any pure equilibrium $(x_1^{**}, \ldots, x_n^{**})$ of this game (a pure equilibrium exists by $A0$), we have $x_n^{**} = BR_n(x_1^{**}, \ldots, x_{n-1}^{**})$ and $x_j^{**} = x_j^*$ for all $j \in \{2, \ldots, n-1\}$. Therefore, player 1's equilibrium payoff

$$\begin{aligned} &u_1(x_1^{**}, x_2^*, \ldots, x_{n-1}^*, BR_n(x_1^{**}, x_2^*, \ldots, x_{n-1}^*)) \\ &> u_1(x_1^*, x_2^*, \ldots, x_{n-1}^*, BR_n(x_1^*, x_2^*, \ldots, x_{n-1}^*)), \end{aligned}$$

a contradiction with x_{-n}^* being a Nash equilibrium of g_n. A similar reasoning applies to players 2 to $n-1$. □

Proof of Theorem 3 The proof is similar to the proof of Theorem 1 (p 200) of Benoit and Krishna (1985). First, suppose that for each player i, there exists a pure Nash equilibrium e^i of g with $u_i(e^i) > \omega_i$. Let $u \in \mathbb{R}^n$ a profile of individually rational and feasible payoffs. From Theorem 1 of Benoit and Krishna, for any $\varepsilon > 0$, there exists a T^* such for for all $T > T^*$, $s^* = (s_i^*, s_{-i}^*)$ is a pure Nash equilibrium of G^T with equilibrium payoff $U = u(s^*)$, and $||U - u|| < \varepsilon$. From Corollary 1, it then follows that there exists a subgame perfect equilibrium of the commitment game with equilibrium payoff U, the desired result.

Second, suppose that for each player i, there exists a self-enforcing commitment ${}^iX^*$, which implements the action profile e^i of g with $u_i(e^i) > \omega_i$. From Theorem 2, e^i is a Nash equilibrium of $g({}^iX^*)$ and, for any game $g(X_j \times {}^iX_{-j}^*)$, there exists a Nash equilibrium of it that gives a lower payoff to player j than $u_j(e^i)$. If we can show that there exists self-enforcing commitments that implement the finite repetition of e^i, then we can replicate the proof of Benoit and Krishna to get the desired result.

Lemma 2 *Let* (X^*, x^*) *an equilibrium outcome of* $\Gamma(g)$*. There exists a Nash equilibrium of* G^T *with* $(R^*, x^*, \ldots, x^*)$ *as its equilibrium path.*

Proof For each player $i \in N$, define $R_i^* \subseteq S_i$ as the set of strategies that satisfies: $\{a_i \in A_i : a_i = s_i(h^0) \text{ for all } s_i \in R_i^*\} = X_i(h^0) = X_i^*$, $X_i(h^t) = X_i^*$ for any history h^t of the form $h^t = (h^0, \underbrace{x^*, ..., x^*}_{t \text{ times}})$, $t \leq T$. Let $R^* = \times_i R_i^*$. We want to show that $(R^*, \underbrace{x^*, ..., x^*}_{T \text{ times}})$ can be sustained as an equilibrium path.

To do so, consider the following strategies of the commitment game $\Gamma(G^T)$.[28] For each player i, define $\gamma_i(d^0) = R_i^*$, $\gamma_i(d^0, R^*) = x_i^*$, $\gamma_i(d^t) = x_i^*$ for all histories d^t of the form $d^t = (R^*, \underbrace{x^*, ..., x^*}_{t \text{ times}})$, $\gamma_i(d^t) = m_i^j$ with m_i^j the action that "min-max" player j in the game g for all histories of the form $d^t = (R^*, \underbrace{x^*, ..., x^*}_{t-1 \text{ times}}, (x_j, x_{-j}^*))$ with $x_j \neq x_j^*$, $\gamma_i(d^t) = m_i^j$ for all histories d^t following the history $d^{t-1} = (R^*, \underbrace{x^*, ..., x^*}_{t-2 \text{ times}}, (x_j, x_{-j}^*))$, and are left unspecified after any other histories following the history $d^1 = (d^0, R^*)$. We still have to define the strategies following a deviation from R_j^* by player j. Let $R_j \neq R_j^*$ be the deviation of player j at the initial history. Define $X_j^1 = \{a_j \in A_j : a_j = s_j(h^0) \text{ for all } s_j \in R_j\}$ and e^1 a Nash equilibrium of the game $g(X_j^1 \times X_{-j}^*)$. We can now define strategies following a deviation by player j as follows: $\gamma_i(d^0, (R_j, R_{-j}^*)) = e^1$, $\gamma_i(d^t) = m_i^j$ for the history $d^t = (d^0, (R_j, R_{-j}^*), e^1)$, $\gamma_i(d^t) = m_i^j$ for all histories of the form $d^t = (d^0, (R_j, R_{-j}^*), e^1, m^j, \ldots, m^j)$, and left unspecified after any other histories following the history $d^1 = (d^0, (R_j, R_{-j}^*))$. Finally, if two players or more deviate from the commitment R^*, strategies are left unspecified.

Before going further, let us explain in words the strategies constructed above. Players commit to choose an action from the set X^* if the profile x^* has been played in previous periods, and do not restrict their set of actions, otherwise. Moreover, if all players commit to R_i^*, the strategy specifies that they play x^* as long as x^* was played in the past and "min-max" a deviating player, otherwise. Finally, if one player deviates from the commitment R_i^*,

[28] Do not confuse with the commitment game $\Gamma(g)$.

the strategy requires players to play first a Nash equilibrium, and then to "min-max" the deviating player for the remaining periods.

Let us show that the strategies define above constitute a Nash equilibrium of the commitment game $\Gamma(G^T)$. First, following the commitment by each player to R_i^*, it is easy to show that $(x^*, \ldots, x^*)$ is the equilibrium outcome of the game $G^T(R^*)$. Indeed, if player i deviates at stage t, his payoff is

$$\frac{t-1}{T}u_i(x^*) + \frac{1}{T}u_i(x_i, x_{-i}^*) + \frac{T-t-1}{T}\overline{\omega}_i,$$

which is smaller than $u_i(x^*)$ since x^* is a Nash equilibrium of $g(X^*)$ and $u_i(x^*) \geq \overline{\omega}_i$ by Corollary 2. Second, following a deviation by player i from R_i^* to R_i, the strategies implies that the payoff to player i is

$$\frac{1}{T}u_i(e^1) + \frac{T-1}{T}\overline{\omega}_i \leq u_i(x^*),$$

where the inequality comes from Theorem 2 and Corollary 2. However, it remains to check that any player $j \neq i$ can play m_j^i after a deviation by player i from e^1. Indeed, if player i deviates from e^1 and unilaterally induces the profile x^*, player i's opponents are then committed to X_{-i}^*, and might not be able to play m_{-i}^i. We show, however, that the profile x^* has to differ from e^1 in at least two components, hence it is not reachable by a unilateral deviation from e^1. Since e^1 is a Nash equilibrium of $g(X_i^1 \times X_{-i}^*)$, we have that $u_i(e_i^1, e_{-i}^1) > u_i(x_i, e_{-i}^1)$ for all $x_i \in X_i$ with a strict inequality since the game is generic. It follows that if x^* is reachable by a unilateral deviation from all Nash equilibria of $g(X_i \times X_{-i}^*)$, we have $u_i(e^1) > u_i(x^*)$ for all equilibria e^1 of $g(X_i \times X_{-i}^*)$, a contradiction with the fact that x^* is implementable in $\Gamma(g)$. □

Lemma 2 implies that there exist self-enforcing commitments that implement the play of a finite sequence of x^*, where x^* is implementable by the commitment X^* in the game $\Gamma(g)$. An immediate extension of Lemma 2 is that there exist self-enforcing commitments that implement any sequence of $({}^1x^*, {}^2x^*, \ldots, {}^nx^*)$, where ${}^ix^*$ is implementable by ${}^iX^*$ in $\Gamma(g)$, and is played P_i periods. To complete the proof, it remains to reproduce the construction of Benoit and Krishna (1985). This is left to the reader. □

Proof of Theorem 4 Let $a^* \in A$ be a profile of actions of g, which satisfies the common threat property, such that $u(a^*)$ is individually rational and satisfies the no-reward condition. Let m^* be the common threat ("min-max"). Let $\tau \in \mathbb{N}$.

For each player i, define the set of strategies $R_i^* \subseteq S_i$ of the game G^T that satisfies: $X_i^*(h^{T-\tau}) = \{a_i^*\}$ if $h^{T-\tau} = (h^0, \underbrace{a^*, \ldots, a^*}_{T-\tau \text{ times}})$, $X_i^*(h^t) = \{a_i^*\}$ for all histories h^t that follows $(h^0, \underbrace{a^*, \ldots, a^*}_{T-\tau \text{ times}})$, $X_i^*(h^t) = \{m_i^*\}$ if $h^t \neq (h^0, \underbrace{a^*, \ldots, a^*}_{t \text{ times}})$, and $X_i^*(h^t) = \{m_i^*\}$ for any subsequent histories.

In words, player i commits to a_i^* for the last τ periods of the game if the profile a^* has always been played in the past, and commits to m_i^* as soon as one player deviates from the profile a^*. Commitments thus resemble commitments to trigger strategies, hence called *trigger commitment*.

We can now construct (part of) the strategy γ_i^* of player i in the game $\Gamma(G^T)$. Let $\gamma_i^*(d^0) = R_i^*$, $\gamma_i^*(d^1) = a_i^*$ if $d^1 = (d^0, X^*)$, $\gamma_i^*(d^t) = a_i^*$ if $d^t = (d^0, R^*, \underbrace{a^*, \ldots, a^*}_{t-1 \text{ times}})$, $\gamma_i^*(d^t) = m_i^*$ if $d^t = (d^0, R^*, \underbrace{a^*, \ldots, (a_j, a_{-j}^*)}_{t-1 \text{ times}})$ with $a_j \neq a_j^*$, and unspecified, otherwise.

Following the history $d^1 = (d^0, X^*)$, the payoff of player i under the strategy profile γ^* is $u_i(x^*)$. If player i deviates at time $t \leq T - \tau$, his maximum payoff is

$$\frac{t-1}{T} u_i(a^*) + \frac{1}{T} \max_{a \in A} u_i(a) + \frac{T-t}{T} \omega_i,$$

which is maximized if $t = T - \tau$. Let τ_i^* be a solution of

$$\tau_i^*(u_i(a^*) - \omega_i) \geq \max_{a \in A} u_i(a) - u_i(a^*).$$

A solution exists since $u_i(a^*) > w_i$. Define $\tau^* = \max_i(\tau_i^*)$. Hence, for $T \geq \tau^*$, γ^* induces a Nash equilibrium in the subgame following the history (d^0, R^*) with equilibrium payoff $u_i(a^*)$ to each player i.

Let us show that γ^* also induces a Nash equilibrium in the subgame following $d^1 = (d^0, (R_i, R_{-i}^*))$ for any $i \in N$, i.e., following any deviation by player i from the commitment R_i^*, that is not profitable.

First, consider the deviating player i. From the previous arguments, player i has a profitable deviation if and only if the profile a^* has been played during the first $T-\tau$ periods following his deviation to R_i. Indeed, if a profile $a \neq a^*$ has been played earlier, this triggers the commitment of player i's opponents to m^*_{-i} for the remaining periods of the game. Hence, if $\tau \geq \tau^*$, this gives to player i a payoff lower than $u_i(a^*)$. The key point to observe is the common threat property allows to "min-max" player i even though he might not be the one to deviate from a^* in the subgame.

Let us now consider player $j \neq i$. We want to show that $(\underbrace{a^*, \dots, a^*}_{T-\tau \text{ times}}, \dots)$ cannot be an equilibrium path of the subgame $G(R_i \times R^*_{-i})$ if a^* satisfies the no-reward condition. Since a^* satisfies the no-reward condition, there exists a player $j \neq i$ and a $\tau^{**} \geq \tau^*$ such that

$$\frac{T-\tau^{**}-1}{T}u_j(a^*) + \frac{1}{T}u_j(a^*) + \frac{1}{T}\sum_{s=1}^{\tau^{**}} u_i(\hat{a}_j(s), a^*_{-j}) <$$
$$\frac{T-\tau^{**}-1}{T}u_j(a^*) + \frac{1}{T}\max_{a_j \in A_j} u_j(a_j, a^*_{-j}) + \frac{\tau^{**}}{T}\min_{a_i \in A_i} u_j(a_i, m^*_{-i}),$$

with $(\hat{a}_i(s))_{s=1}^{\tau^{**}}$ solution of

$$\max_{\{a_i(s)\} \in A_i^{\tau^{**}}} \sum_{s=1}^{\tau^{**}} u_j(a_i(s), a^*_{-i}) \tag{4}$$

subject to

$$\sum_{s=1}^{\tau^{**}} u_i(a_i(s), a^*_{-i}) > \tau^{**} u_i(a^*).$$

Therefore, since a^* satisfies the no-reward condition, there exists a player $j \neq i$ for whom it is profitable to deviate from a^* at period $T-\tau^{**}$.

To conclude, if G^T is supermodular game, $G^T(R_i \times R^*_{-i})$ also is, and therefore has a pure Nash equilibrium. From the above arguments, the payoff to player i in any pure Nash equilibrium of $G^T(R_i \times R^*_{-i})$ is smaller than $u_i(a^*)$, hence the deviation is not profitable. Finally, suppose that G^T is a two-player game. If player i is indifferent between any strategy that prescribes to play a^*_i for the first $T-\tau^{**}$ periods and another strategy, then it is payoff is at most $u_i(a^*)$ from the above arguments. And if player i plays with probability one a strategy that prescribes to play a_i^* for the first $T - \tau^{**}$ periods, player j cannot be indifferent between a strategy that prescribes to play a_j^* for the first $T - \tau^{**}$ and a strategy that prescribes to play $a_j \neq a_j^*$ at period $T - \tau^{**}$. □

References

[1] S. Bade, G. Haeringer, L. Renou, Bilateral commitment, University of Leicester, 2007.

[2] J.P. Benoît, V. Krishna, Nash Equilibria of Finitely Repeated Games, *International Journal of Game Theory* 16 (1987), 197-204.

[3] C-F. Chou, J. Geanakoplos, The Power of Commitment, Cowles Foundation Discussion Paper No. 885, 1988.

[4] E. van Damme, S. Hurkens, Commitment Robust Equilibria and Endogenous Timing, *Games and Economic Behavior* 15 (1996), 290-311.

[5] F. Echenique, Extensive-form Games and Strategic Complementarities, *Games and Economic Behavior* 46 (2004), 348-364.

[6] A. Faña-Medín, I. García-Jurado, J. Méndez-Naya, L. Méndez-Naya, Unilateral Commitments in the Finitely Repeated Prisoners Dilemma, *Annals of Operations Research* 84 (1998), 187-194.

[7] I. García-Jurado, J. González-Díaz, The Role of Commitment in Repeated Games, *Optimization*, 55 (2006), 1-13.

[8] I. Gilboa and E. Zemel, Nash and Correlated Equilibria: Some Complexity Considerations, *Games and Economic Behavior* 1 (1989), 80-93.

[9] J. González-Díaz, Finitely Repeated Games: A Generalized Nash folk Theorem, *Games and Economic Behavior*, 55 (2006), 100-111.

[10] J. H. Hamilton, S. M. Slutsky, Endogeneous Timing in Duopoly Games: Stackelberg or Cournot equilibria, *Games and Economic Behavior* 2 (1990), 29-46.

[11] ——, Endogenizing the Order of Moves in Matrix Games, *Theory and Decision* 34 (1993), 47-62.

[12] M. Jackson, S. Wilkie, Endogenous Games and Mechanisms: Side Payments among Players, *Review of Economic Studies* 72 (2005), 543-566.

[13] A.T. Kalai, E. Kalai, E. Lehrer, D. Samet, Voluntary Commitments Lead to Efficiency, Mimeo, Tel Aviv University, http://www.math.tau.ac.il/ lehrer/Papers/.

[14] R.D. McKelvey, A. McLennan, Computation of Equilibria in Finite Games, in *Handbook of Computational Economics* Eds. H.M. Amman, D.A. Kendrick, J.Rust, Elsevier, 1996.

[15] R.D. McKelvey, A. McLennan, T. Turocy, Gambit: Software Tools for Game Theory, Version 0.2007.01.30, http://econweb.tamu.edu/gambit.

[16] P. Milgrom, J. Roberts, Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities, *Econometrica* 58 (1990), 1255-1277.

[17] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*, 1994, MIT Press, Cambridge, Massachusetts.

[18] R. Romano, H. Yildirim, On the Endogeneity of Cournot-Nash and Stackelberg equilibria: Games of Accumulation, *Journal of Economic Theory* 120 (2005), 73-107.

[19] T. Schelling, *The Strategy of Conflict*, 1960, Harvard University Press, Cambridge, Massachusetts.

[20] D.M. Topkis, *Supermodularity and Complementarity*, 1998, MIT Press, Cambridge, Massachusetts.