University of Leicester

DEPARTMENT OF ECONOMICS

BILATERAL COMMITMENT

Sophie Bade, Penn State University, USA
Guillaume Haeringer, Universitat Autònoma de Barcelona, Spain
Ludovic Renou, University of Leicester, UK

Working Paper No. 08/20
June 2008

Bilateral Commitment*

Sophie Bade[†] Guillaume Haeringer[‡] Ludovic Renou[§]

June 25, 2008

Abstract

We consider non-cooperative environments in which two players have the power to commit but cannot sign binding agreements. We show that by committing to a set of actions rather than to a single action, players can implement a wide range of action profiles. We give a complete characterization of implementable profiles and provide a simple method to find them. Profiles implementable by bilateral commitments are shown to be generically inefficient. Surprisingly, allowing for gradualism (i.e., step by step commitment) does not change the set of implementable profiles.

Keywords: Commitment, self-enforcing, generic inefficiency, agreements, Pareto-improvement.

JEL Classification Numbers: C70, C72, H87.

*We are grateful to Dirk Bergemann, Matthew O. Jackson, Doh-Shin Jeon, Jordi Massó, Claudio Mezzetti, Efe Ok, Andrés Rodríguez-Clare, Ben Zissimos, seminar participants at Adelaide, Auckland, Barcelona JOCS, Bocconi, Leicester, Maastricht, Paris, University of New South Wales, Sydney, and Vienna for helpful comments and suggestions. We would like to warmly thank John van der Hoek and Guillaume Carlier for patiently and brilliantly answering our questions on differential topology and the theory of non-smooth optimizations. Haeringer acknowledges a financial support from the Spanish Ministry of Science and Technology through grants BEC2002-02130 and SEJ2005-01481/ECON, and from the Barcelona Economics - XREA program. Renou acknowledges financial support from the Faculty of the Professions.

†Department of Economics, Penn State University, 518 Kern Graduate Building, University Park, PA 16802-3306, USA. Phone: (814) 865 8871, Fax: (814) 863 - 4775 sub18@psu.edu

‡Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain. Phone: (+34) 93 581 12 15, Fax: (+34) 93 581 20 12. Guillaume.Haeringer@uab.es

§Department of Economics, University of Leicester, LE1 7RH Leicester, UK Tel: ++44 (0) 116 252 3980, lr78@le.ac.uk

1 Introduction

An essential insight of Schelling (1960) is that a player can strengthen his position by committing to some actions. For example, a monopolist can deter entry by committing to fight any eventual entry. However, suppose that the potential entrant can simultaneously commit to enter the market, say by installing capacities. The monopolist commitment is then inconsistent with the entrant commitment: the monopolist would rather accommodate the entry. What are the consistent commitments in such a game? The main theme of this paper is precisely to study situations, as the above example, in which all the players in a game can simultaneously and perfectly commit to *subsets* of actions before playing a game. Before going further, we like to stress that players are not assumed to commit to a particular action, but rather to rule out sets of actions.[1] Classical examples of such commitments are firms choosing capacity constraints, an army general burning a bridge behind his troops, a candidate promising not to raise taxes by more than 5%, or a seller publicly announcing a menu of tariffs. In all these cases, reneging on one's commitment is either physically impossible or too costly to be considered.[2]

To model the possibility of commitment in games, we embed a strategic-form game G into a two-stage game, in which players can restrict their action spaces in the first stage (the commitment stage), and play the game induced by their commitments in the second stage. Payoffs are determined as in the original game G. We call this two-stage game, *a game of commitment*, and consider the subgame perfect equilibria of games of commitment. More precisely, we are interested in the complete characterization of the action profiles of G that are implementable by commitments, that is, the action profiles played in the second stage in any subgame perfect equilibrium. Our main contribution is to provide a complete characterization of the implementable profiles of actions for two-player games with closed real intervals as action spaces, strictly quasi-concave payoff functions, and commitments to closed subintervals of the original action space. These assumptions are met by many economic models, e.g., Cournot and differentiated Bertrand duopoly games, games of tax competition, etc. In equilibrium, commitments are *self-enforcing* in the sense that they are sustained by a simple sequential game structure, without assuming any punishment scheme against deviating players.

The question whether an action profile is implementable by a commitment is a complex one. To see this, note that any action profile belongs to an infinite set of restricted action spaces.

[1] Our approach to commitment is shared by Hart and Moore (2004), who study the case of two contracting parties who can restrict the set of outcomes over which they will bargain. One of the main differences between their work and ours is that they assume that some uncertainty is being resolved after players have committed to a set of outcomes and before the parties bargain over the final outcome. Without such uncertainty, parties would fully commit in the first period in the framework of Hart and Moore (2004).

[2] See Caruana and Einav (2005) for a model in which commitment arises endogenously.

Therefore, to find out whether a profile x is implementable by a commitment, we would have to check whether it is implementable by any one of these infinitely many pairs of restricted action spaces. The main result of this paper is that an action profile is implementable *if and only if* it is implementable by what we call a "*simple commitment*." In a simple commitment, one player (he) commits to a single action, and the other player (she) truncates her action space at either the top or the bottom. Moreover, the truncation is at her (original) best-reply to the single action her opponent is committed to. It follows that for any action profile, there are only *four* such simple commitments. This result drastically reduces the complexity of our problem.

To get more intuition on the profiles of actions that are implementable, note that all Nash equilibria of the original game are implementable. The intuition for this result is simple. Suppose that each player commits to his equilibrium action in the first stage of the game of commitment, and plays a Nash equilibrium in any induced game. Given the commitment of a player, the other player has obviously no incentive to deviate as he is already playing the best-reply to the single action in the commitment set of his opponent. Similarly, all 'lead-follow' outcomes are implementable. A 'lead-follow' outcome is a subgame-perfect equilibrium outcome of the sequential version of G, in which one player is moving first and the other follows (Stackelberg outcomes in duopoly games). To implement such outcomes it suffices that the 'leader' commits to a single action (his action in the lead-follow profile) and the other player does not restrict his action space at all. This is not accidental, we show that all action profiles that can be implemented by a game of commitment can be described as the equilibrium outcome of a generalized sequential version of the game under consideration. Important insights about following and leading in sequential games apply to the game of strategic commitment. We use these insights to translate our characterization results into a geometrical representation. We can show in particular that with a further restriction to games with strategic complementarities the best reply curves alone suffice to characterize all implementable profiles, in this case the set of implementable profiles is bounded by the Nash- and follow-lead equilibrium outcomes.

We pursue our characterization by considering a variant of our commitment game, allowing players to commit in several steps. In a recent paper, Lockwood and Thomas (2002) indeed show that gradualism may enforce partial cooperation that is not attainable in one step commitment. It turns out that this is not the case in our setup: a profile is implementable in T rounds of commitment *if and only if* it is implementable in one round.

Finally, an important question is whether bilateral commitment may help players to improve over the status quo, i.e., the Nash equilibria of the original game. First, we show that the players cannot, generically speaking, implement efficient outcomes using commitments.[3] Second, we

[3]This result parallels Dubey's (1986) theorem that shows that Nash equilibria of smooth games are generically inefficient.

show that when 'lead-follow' equilibria, which are always implementable by commitment, do not give both players a higher payoff than the Nash equilibria, then no Pareto improvements are implementable in the important class of games with strategic complementarities and constant consonance, that is, when the payoff of a player is monotone in the action of his opponent. Finally, we give an example of a game with a non-monotonic best reply curve in which parties can Pareto improve upon a unique Nash equilibrium even though the 'follow-lead' equilibria do not Pareto dominate the Nash equilibrium. Thus, we conclude on a positive note: bilateral commitments might improve the welfare of each player.

The idea that the power to commit oneself can be beneficial has received a great deal of attention in economics. A (very) partial list of contributions includes applications in industrial organization (e.g., Dixit (1980) or Spulber (1981)), international trade (e.g., Brander and Spencer (1985)), political economy (e.g., Yildirim (2005)), to name just a few.[4] Most of these applications can be seen as special cases of our theory, in the sense that commitments made in an initial stage restrict the set of actions available in a later stage. Closely related to our work is the literature on endogenous timing in games e.g., Hamilton and Slustky (1990), Amir and Grilo (1999), van Damme and Hurkens (1999), or Romano and Yildirim (2005).[5] The aim of this literature is to obtain Cournot-Nash and Stackelberg outcomes 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, in our model, commitments are not restricted to commitments to single actions.[6] In other words, a commitment in our game might leave something open to change. Second, our purpose is *not* to endogenize the timing of moves in games, but to explore how the ability to commit affects the equilibrium payoffs, and its welfare properties. Thus, our approach is conceptually different from the approach followed in the endogenous timing literature. Our results, however, parallel the results in this literature insofar as the additional flexibility in the choice of commitments we postulate yields a range of implementable profiles that is — in a sense to be defined more precisely — bounded by the *Cournot-Nash* and *Stackelberg* outcomes as extreme cases. Romano and Yildirim's (2005) paper is the closest to ours, though our work covers a much wider range of cases. First, our commitment technology is more general in that players can restrict their action spaces from the bottom *and* the top. Second, we do not assume differentiability of the payoff functions, monotonic best replies, a unique interior Nash equilibrium of the original game; assumptions all made by Romano and Yildirim. Third, like us, Romano and Yildirim extend their commitment

[4]Kreps and Scheinkman (1983) differs from our work in that capacity commitment in the first stage does not affect the action set in the second stage, but the payoffs.

[5]See also Saloner (1987), Gale (2001) or Henkel (2002) among others.

[6]A notable exception is Romano and Yildirim (2005), who consider commitments to lower or upper bounds, but not both.

game by allowing players to commit in several steps. In this context, they also show that if a profile is implementable in T rounds of commitment ($T \geq 2$), then it is also implementable in 1 round of commitment, but admit that they are unable to prove the converse, which we do; thereby improving upon their results. Lastly, we also improve upon their work by extensively discussing the welfare implications of our model and providing a geometric characterization of the implementable profiles. Thus, our contribution is both conceptual and technical. In some sense, our paper answers the question: when can we assume without loss of generality that a player commits to a single action and the other player truncates his action space only at the top or bottom? We also note that in Cournot duopoly games, the set of implementable actions in our commitment games is equivalent to the one in Romano and Yildirim, thus rationalizing their assumption of commitment to lower bounds only. However, in other examples e.g., rent seeking games, their assumption is not without loss of generality.

We should also mention Jackson and Wilkie (2005). They also allow players to modify the game to be played in a pre-play stage. The main difference between their work and ours lies in the set of permissible modifications. While Jackson and Wilkie (2005) allow players to commit to utility transfers in the second period, we allow players to discard any subset of actions in the pre-play stage. These different pre-play modifications yield different results. Nash equilibria can always be implemented in our framework but need not be implementable in theirs. On the other hand, they show, like us, that pre-play modifications do not necessarily make efficient outcomes implementable. Finally, Renou (2006) provides a complete characterization of the equilibrium payoffs in commitment games induced by n-player finite games.

This paper is organized as follows. In Section 2, we give a detailed description of the environment faced by the players, and define what we call the game of commitment. Section 3 presents some preliminary results. In Section 4, we completely characterize the set of action profiles that are implementable by self-enforcing bilateral commitment. Section 5 analyzes the welfare implications of self-enforcing bilateral commitment. Section 6 discusses possible extensions. Most proofs are relegated in the Appendix.

2 Games of commitment

2.1 Preliminaries

The initial situation we consider is a two-player strategic-form game $G := \langle N, (Y_i, u_i)_{i \in N} \rangle$ with $N = \{1, 2\}$ the set of players, Y_i the set of actions available to player i, and $u_i : Y_1 \times Y_2 \to \mathbb{R}$ the payoff function of player i. Denote $Y := Y_1 \times Y_2$. We call the opponent of player i, player j. We assume that for each player $i \in \{1, 2\}$, Y_i is a non-empty, compact, convex subset of the real line. Without loss of generality, we take $Y_i = [0, 1]$, for $i \in \{1, 2\}$. For each player i, the payoff function u_i is assumed to be continuous in all its arguments and strictly quasi-concave in y_i, i.e., for all $y_j \in [0,1]$, $y_i \in [0,1]$, $y_i' \in [0,1]$, $y_i \neq y_i'$, and $\alpha \in (0,1)$, $u_i(\alpha y_i+(1-\alpha)y_i', y_j) > \min\{u_i(y_i,y_j), u_i(y_i',y_j)\}$.[7] These assumptions are met by many economic models.

We furthermore assume that players have the ability to unilaterally commit not to play some actions, i.e., to restrict their action sets. Such commitments are assumed to be perfectly binding, meaning that if player i restricts his action set to X_i, any action chosen later on must belong to X_i.

Definition 1 A *(bilateral) commitment* is a pair (X_1, X_2) where for both $i \in \{1,2\}$, X_i is a non-empty, compact and convex subset of $[0,1]$.

Thus, our definition of a commitment imposes on each player a *restriction* of his action space.[8]

Henceforth, we write the restricted action space X_i of player i as a closed real interval $[\underline{x}_i, \overline{x}_i] \subseteq [0,1]$, where $\underline{x}_i$ ($\overline{x}_i$) refers to the minimum (maximum) of player i's restricted action space. Note that player i can also commit to a singleton, in which case $\underline{x}_i = \overline{x}_i$.

It is important to note that a commitment does not necessarily prescribe the choice of an action. In the words of Hart and Moore (2004), "in a bilateral commitment, the players commit not to consider actions not on the list (X_1, X_2), i.e., these actions are ruled out. Ex-post, the players are free to choose from the list of actions specified in the commitment i.e., actions are not ruled in."

We say that the bilateral commitment (X_1, X_2) *induces* the game $G(X) := \langle N, (X_i, u_i^X)\rangle$, where $X = X_1 \times X_2$, and for $i \in \{1,2\}$, $u_i^X(x) = u_i(x)$ for all $x \in X$. Abusing notation, we will drop the superscript X in the sequel. The *induced* game $G(X)$ is thus obtained from the game G by restricting the action sets of the players. We shall use the term '*mother*' to make reference to the original game G. For instance, we shall use the expressions *mother game*, *mother best-reply*, *mother action set*, etc. Similarly, the term '*induced*' will refer to the best reply, action sets etc. in $G(X)$. We denote by $\mathcal{Y}_i$ the collection of all non-empty, compact, convex subsets of $[0,1]$, and

[7]In the words of Moulin (1984), G is a two-player 'nice game.' It is worth noting that the mixed extensions of any finite games do not satisfy our assumptions. First, payoff functions are not strictly quasi-concave in such games. Second, unless the finite game has only two actions per player, mixed action spaces are not a subset of the real line. Consequently, the theory developed in this paper cannot be applied to mixed extensions of finite games.

[8]That restrictions are assumed to be convex subsets is not without loss of generality. In particular it ensures that the game played once players have chosen their restrictions has a Nash equilibrium. Imposing some Lipschitz conditions is sufficient, however, to deal with non-convex restrictions. We also note that imposing convex strategy sets is a common assumption in the economic literature.

define $\mathcal{Y} := \prod_{i\in\{1,2\}} \mathcal{Y}_i$.

2.2 Games of commitment

Given the strategic-form game G, the *game of commitment* $\Gamma(G)$ is a two-stage game with almost perfect information, in which:

Stage 1. Both players simultaneously choose action sets $X_i \in \mathcal{Y}_i$.

Stage 2. Players play the induced strategic form game $G(X)$.

A *strategy* for a player i in the game $\Gamma(G)$ (for short, Γ), is a pair $s_i = (X_i, \sigma_i)$ where $X_i \in \mathcal{Y}_i$, and σ_i is a mapping from $\mathcal{Y}$ to $[0,1]$ such that $\sigma_i(X) \in X_i$, for all $X \in \mathcal{Y}$. That is, a strategy for a player prescribes a choice of a restriction X_i (first-stage action) and, for each possible choice of a restriction for *both* players in the first-stage, an action $x_i \in X_i$ (second-stage action). The *outcome* of a strategy profile $s = (s_i)_{i\in\{1,2\}}$ is the pair (X, x) where $x_i = \sigma_i(X)$ for each player $i \in \{1,2\}$. The payoffs over outcomes (X, x) are assumed to only depend on the action profiles chosen in the second stage of the game and are given by the payoffs of the induced game $G(X)$. That is, we assume that player i derives utility $u_i(x)$ from outcome (X, x). If (X, x) is the outcome of strategy profile s we call x the *result* of s.

The central concept of this paper is the concept of implementation by commitment, which we now define.

Definition 2 An action profile x is *implementable by commitment* X if the pair (X, x) is the outcome of a subgame-perfect equilibrium of Γ.

Hence, a profile x is implementable by commitment if it is a (stage 2) result of a subgame-perfect equilibrium of Γ. In this paper, we focus on subgame-perfect equilibria in pure strategies.

3 Games induced by commitments

We first derive some results concerning the proper subgames of Γ, namely the set of all induced games $G(X)$. The proofs of the results presented below, Lemmata 1 and 2 are in our companion paper, Bade, Haeringer and Renou (2005).

Define $BR_i : [0,1] \to [0,1]$, the *(mother) best-reply* of player i in the game G, with for $y_j \in [0,1]$,

$$BR_i(y_j) = \{y_i \in [0,1] : u_i(y_i, y_j) \geq u_i(y_i', y_j) \text{ for all } y_i' \in [0,1]\}.$$

When players commit to play in the set X, the best-reply map $br_i^X : X_j \to X_i$ of player i is defined similarly, bearing in mind that now player i cannot choose an action outside X_i, that is, for all $x_j \in X_j$,

$$br_i^X(x_j) = \{x_i \in X_i : u_i(x_i, x_j) \geq u_i(x_i', x_j) \text{ for all } x_i' \in X_i\}.$$

We will denote the best-reply map $br_i^{X_i \times [0,1]}$ by $br_i^{X_i}$. That is, $br_i^{X_i}$ is the restricted best-reply of player i when he is committed to X_i and player j can choose any action in $[0, 1]$. Note that best-reply maps are non-empty, single valued and continuous. Furthermore, the strict quasi-concavity of payoff functions enables us to easily characterize the mapping br_i^X as a function of BR_i and X.

Lemma 1 *Player i's best-reply function in $G(X)$, $br_i^X : X_j \to X_i$, is*

$$br_i^X(x_j) = \begin{cases} \underline{x}_i & \text{if } BR_i(x_j) < \underline{x}_i\,, \\ BR_i(x_j) & \text{if } \underline{x}_i \leq BR_i(x_j) \leq \overline{x}_i\,, \\ \overline{x}_i & \text{if } \overline{x}_i < BR_i(x_j)\,. \end{cases}$$

In words, the best-reply map br_i^X of the restricted game $G(X)$ agrees with the best-reply map BR_i of the mother game G on the set $\{x_j \in X_j : BR_i(x_j) \in X_i\}$, and is either $\underline{x}_i$ or $\overline{x}_i$, otherwise. Lemma 1 is illustrated in Figures (1a) and (1b). In the former it displays a mother best-reply of player j and in the latter the restricted best-reply when he commits to $[\underline{x}_j, \overline{x}_j]$.

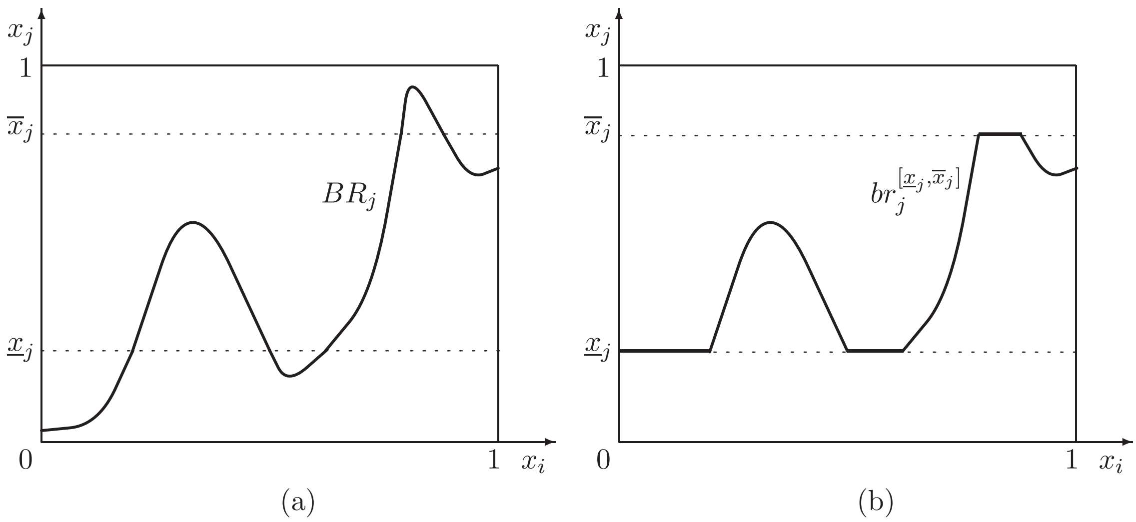

Figure 1: Mother and restricted best-replies

Denote $N(G)$ and $N(G(X))$ the set of Nash equilibria of G and $G(X)$, respectively. Observe that the mother game G as well as any induced game $G(X)$ has a Nash equilibrium in pure actions. Our next lemma states that if a profile of actions x^* is an equilibrium of $G(X)$, but is not an equilibrium of the mother game G, then $x^* \in \mathrm{bd}_Y(X)$, the relative boundary of X in Y.[9]

Lemma 2 *If $x^* \in N(G(X)) \setminus N(G)$, then $x^* \in \mathrm{bd}_Y(X)$.*

Lemma 2 states that if a commitment X^* implements a result x^* that is not an equilibrium of G, then it must be the case that for at least one player, say i, the action x_i^* is either the maximum or the minimum of X_i^*. Lemma 2 thus provides a first intuition about the set of implementable profiles. Namely, if the implemented profile is not a Nash equilibrium of the mother game G, then the action of at least one player identifies with the boundary of his restricted action space.

4 Implementation by commitments

4.1 Existence

We start by observing that the existence of a subgame-perfect equilibrium of Γ is not, *a priori*, guaranteed, for the cardinality of each player's strategy set in Γ is uncountable. It turns out, however, that the issue of equilibrium existence in our case is easily solved.[10]

Proposition 1 *The game of commitment has an equilibrium.*

Proof. Since $\Gamma(G)$ is a finite horizon game, we can use the one-shot deviation property to check that a profile is an equilibrium —see Osborne and Rubinstein (1994, p. 103). Choose $y^* \in N(G)$ and consider for each player i the strategy $s_i^* = (\{y_i^*\}, \sigma_i^*)$, with $(\sigma_i^*(X))_{i\in\{1,2\}}$ a Nash equilibrium of $G(X)$ for any first-stage actions (commitment) X. By construction, no player can profitably change his second-stage action. Observe that since for both $i \in \{1, 2\}$ we have $y_i^* = BR_i(y_j^*)$, neither player can obtain a strictly higher payoff than $u_i(y^*)$. Therefore, given the restriction of player i to $\{y_i^*\}$, player j cannot increase his utility by changing his restriction on his action space. ■

The key observation in the proof of Proposition 1 is that any Nash equilibrium of the mother game G is implementable. So, commitments have the power to perpetuate an existing situa-

[9]Let (Y, d) be a metric space and $X \subset Y$. A point x is a boundary point of X in Y if each open neighborhood U of x satisfies $U \cap X \neq \emptyset$ and $U \cap (Y \setminus X) \neq \emptyset$. The set of all boundary points of X in Y is $\mathrm{bd}_Y X$. For instance, if $Y = [0, 1]$, $\mathrm{bd}_Y[0, 1/2] = \{1/2\}$ while $\mathrm{bd}_Y[1/3, 2/3] = \{1/3, 2/3\}$.

[10]See, for instance, Harris *et al.* (1995) for results on the existence of subgame-perfect equilibria for continuous games with almost perfect information. It is worth noting that Proposition 1 holds independently of the number of players involved in the mother game G.

tion.[11] Moreover, it should be noted that uniqueness is clearly not guaranteed. For instance, if G has a multiplicity of equilibria, then we can already construct a multiplicity of subgame-perfect equilibria of Γ.

4.2 A complete characterization

We are now ready to characterize the set of all action profiles that can be implemented by a commitment. The main result of this section is that if a profile of actions x is implementable, then it is implementable by one of a very small number of bilateral commitments, those that we call *simple*.

Definition 3 A bilateral commitment X is *simple* if it has the form $(\{x_i\}, [0, BR_j(x_i)])$ or $(\{x_i\}, [BR_j(x_i), 1])$.

In a simple commitment, one player takes an extreme position, that of excluding all but one action. The other player, player j, truncates his action space either from below or from above, but not both. Moreover, the truncation is at his best-reply to the only action in player i's extreme commitment. We are now ready to formally state the main result of this section:

Theorem 1 *An action profile x^* is implementable by a bilateral commitment if and only if it is implementable by a simple bilateral commitment.*

Before proving this characterization result, let us briefly comment on the implications of this theorem (see Section 5.5. for more on this). If we want to check whether a particular profile can be implemented by a commitment, we only need to check whether it can be implemented by a simple commitment. This is a very manageable task, as for any action profile x^*, there are exactly 4 simple commitments that could implement it. These commitments are:

$$([0, BR_1(x_2^*)], \{x_2^*\}), \quad ([BR_1(x_2^*), 1], \{x_2^*\}),$$
$$(\{x_1^*\}, [0, BR_2(x_1^*)]), \quad (\{x_1^*\}, [BR_2(x_1^*), 1]) .$$

It is not difficult to check whether an action profile can be implemented by one of these four simple commitments. Indeed, to check whether x^* is implementable by $(\{x_1^*\}, [0, BR_2(x_1^*)])$, it suffices to check whether player 1 has an incentive to change his restricted action space. Observe that in the second stage, neither player has an incentive to deviate (player 2 will be playing the

[11] In a related paper, Jackson and Wilkie (2005) propose a model in which players can commit to utility transfers conditional on actions being played. They notably show that Nash equilibria of the game without transfer, the mother game, might not be implementable, while they are in our paper. An essential difference between their paper and our paper is that commitments can be undone in their paper by transferring back, while it is not possible in our paper.

mother best-reply to player 1's action, and player 1 does not have any choice). Furthermore, given that player 1 commits to $\{x_1^*\}$, player 2 does not have an incentive to alter his commitment, the mother best-reply to x_1^* is already contained in $[0, BR_2(x_1^*)]$). Therefore, we only need to check whether player 1 has an incentive to deviate in the first stage of the game. Notice that for any restriction X_1 player 1 may choose the profile played in the second stage must be a Nash equilibrium of $G(X_1 \times X_2^*)$. So, if player 1 chooses the restriction $\{x_1\}$ for some $x_1 \in [0,1]$, the second stage result will be $(x_1, br_2^{[0,BR_2(x_1^*)]}(x_1))$. Consequently, the action profile x^* is an equilibrium if x_1^* solves the following optimization program:

$$\max_{x_1 \in [0,1]} \quad u_1(x_1, br_2^{[0,BR_2(x_1^*)]}(x_1)) \,. \tag{1}$$

In Section 5.5, we take this optimization program as a starting point for a geometric characterization of implementable profiles.

4.3 Proof of Theorem 1

In this section, we present the main steps leading to Theorem 1 and give intuitions for these intermediate results. Detailed proofs can be found in the Appendix. We start by showing a key result, namely if a result x^* is implementable, then for at least one player $i \in \{1,2\}$, x_i^* is a mother best-reply to x_j^*.

Proposition 2 *Let x^* be implementable by some bilateral commitment X^*. Then $x_i^* = BR_i(x_j^*)$ for at least one player $i \in \{1,2\}$.*

To see the intuition behind Proposition 2, suppose that a profile x^* is implementable by the bilateral commitment X^* such that neither player is using his mother best-reply. From Lemma 2 this means that for both players the constraints imposed by the commitment bind. The continuity of the best replies implies that for all of player 2's actions in a sufficiently small interval $(x_2^* - \varepsilon, x_2^* + \varepsilon)$ around x_2^*, player 1's restricted best reply remains x_1^*. Let us now consider a different restriction for player 2. Take a $\{x_2'\}$ such that x_2' is 1) closer to player 2's mother best-reply to x_1^*, $BR_2(x_1^*)$, and 2) inside the interval $(x_2^* - \varepsilon, x_2^* + \varepsilon)$. (See Figure 2.) The strict quasi-concavity of player 2's payoff function implies that the result (x_1^*, x_2') is strictly preferred to x^*. This implies that player 2 has a profitable deviation, a contradiction with our assumption that x^* is implementable with the bilateral commitment X^*.

Proposition 3 *Let x^* be implementable by some bilateral commitment X^* with $x_j^* = BR_j(x_i^*)$. Then x^* is also implementable by the bilateral commitment X', such that $X_i' = \{x_i^*\}$ and $X_j' = X_j^*$.*

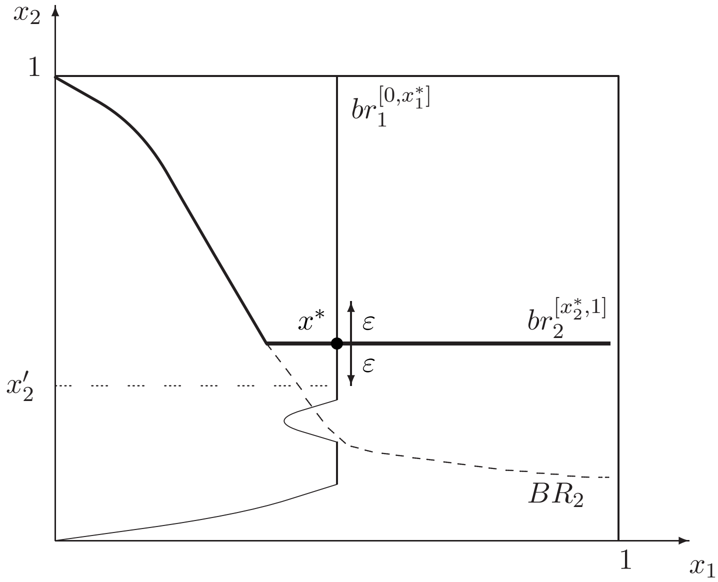

Figure 2: Illustration of Proposition 2

There is a tight connection between Proposition 2 and Proposition 3. By Proposition 2, we know that in any equilibrium outcome (X^*, x^*) of Γ, $x_j^* = BR_j(x_i^*)$ for at least one player $j \in \{1, 2\}$. Imagine now that player i commits to the singleton $\{x_i^*\}$. Since player j can still play $BR_j(x_i^*)$ in the second stage and there player i has no other choice but playing x_i^* in the second stage, player j has no incentive to deviate. If player i can profitably deviate when choosing the restriction $\{x_i^*\}$, he can also profitably deviate when choosing the restriction X_i^*. This, however, cannot be true as we started out with the assumption the (X^*, x^*) is an equilibrium outcome of the game.

The main insight of Proposition 3 is that if $(x_i^*, BR_j(x_i^*))$ is implementable by a bilateral commitment X^*, then it is also implementable by the commitment

$$X' = (\{x_i^*\}, X_j^*). \tag{2}$$

To obtain Theorem 1, it suffices then to show that X_j^* can be reduced to be either $[0, x_j^*]$ or $[x_j^*, 1]$. We establish precisely that in the following proposition.

Proposition 4 *Let x^* be implementable by some bilateral commitment $(\{x_i^*\}, X_j^*)$ with $x_j^* = BR_j(x_i^*)$. Then x^* is also implementable by a commitment X' such that $X_i' = \{x_i^*\}$ and either $X_j' = [BR_j(x_i^*), 1]$ or $X_j' = [0, BR_j(x_i^*)]$.*

Now to prove Theorem 1, take any implementable action profiles x^* and let X^* be a bilateral commitment that implements it. By Proposition 3, we know that the commitment $(\{x_i^*\}, X_j^*)$ for $i \in \{1, 2\}$ does also implement x^*. Finally, from Proposition 4, we know that an action profile that can be implemented by such a commitment can also be implemented by a simple commitment. In sum, these arguments imply that an action profile can be implemented by a commitment only if it can be implemented by a simple commitment. Conversely, any action profile that can be implemented by a simple commitment can be implemented by a commitment. This completes the proof of Theorem 1.

4.4 Multi-period games of commitment

It is often conjectured that the lack of enforcement options may be overcome by considering gradual commitments, thus allowing to implement outcomes that could not be attainable if players can only commit once.[12] The intuition that drives this conjecture is that in a dynamic setting players may find it profitable to make 'small' commitment. Such small commitments might incentive the opponent to also commit but have the merit to minimize the loss if the opponent does not commit. Two central contributions on this issue are Admati and Perry (1991) and Lockwood and Thomas (2002). Admati and Perry (1991) consider a model in which players can make repeated voluntary contributions to finance a project. This latter is implemented only if the sum of the contribution passes a threshold. The game stops as soon as the project is implemented. Lockwood and Thomas (2002) consider a finitely repeated prisoners' dilemma with continuous action space in which at each stage players can only increase their level of cooperation. Both models show that efficient, or nearly efficient outcomes can be obtained.[13] In this section, we follow this line of research by considering a multi-period game of commitment, denoted Γ^T.

In the game Γ^T, players face T periods of commitment and one final stage in which they play the game induced by their commitments. In each period $t = 1, \ldots, T$, players simultaneously restrict their action spaces with the constraint that the restriction at stage t has to be a non-empty, compact, convex subset of the restricted action space at period $t-1$. That is, if X_i^t denotes the restriction of player i at period t then $X_i^{t+1} \subseteq X_i^t$. Finally, in period $T+1$, players play the game induced by the commitment of period T, the game $G(X^T)$.

One may imagine that allowing for several stages of commitment may change the set of

[12]See Schelling (1956) for an early account on this issue.

[13]The models of Admati and Perry (1991) and Lockwood and Thomas (2002) do not separate as clearly as we do the commitment decision from the decision of choosing which action to play. Their models are simply repeated games in which the assumption that at each stage players cannot use an action 'lower' than their action at the previous stage. First, this implies that in their models players can only restrict their action sets by choosing a lower bound (the contribution level in Admati and Perry (1991) or the cooperation level in Lockwood and Thomas (2002)). Second, a key difference is that in their model, the payoff is dependent on the sequence of commitments (lower bounds), while in our model we do assume that commitments do not enter directly the payoff functions.

implementable profiles. In fact, it turns out that in our context this is not the case.

Theorem 2 *For any T a profile of actions x^* is implementable in the multi-period game of commitment $\Gamma^T(G)$ if and only if it is implementable in a game of commitment $\Gamma(G)$.*

The proof of this theorem heavily rests on a result similar to that of Proposition 2, i.e., if x^* is implementable in T rounds of commitment then at least one player is best-replying. A key observation to prove Theorem 2 is that for any equilibrium s^* of Γ^T, we can always construct a new equilibrium profile $\hat{s}$ in which players' first stage restrictions are the same as their last restrictions under s^* (on the equilibrium path), and at all other subsequent stages players do not further restrict their action spaces. Hence, from the perspective of characterizing the set of implementable profiles repeating the number of stages at which players can restrict their action spaces does not enrich our model.

4.5 The geometry of implementable profiles

As already pointed out, Theorem 1 has remarkable implications for the characterization of the implementable action profiles of a game of commitment. To check whether a profile of actions x is implementable, it suffices to follow a simple four-step procedure:

Step 1. Check whether x lies on the graph of the best-reply map of at least one player. If not, then x is not implementable. If yes, go to step 2.

Step 2. Check whether x lies on the best-reply graphs of both players. If yes, then x is implementable since it is an equilibrium of the mother game G. If not, go to step 3.

Step 3. Without loss of generality, assume that $x_j = BR_j(x_i)$. Construct the simple commitments $(\{x_i\}, [0, BR_j(x_i)])$ and $(\{x_i\}, [BR_j(x_i), 1])$. Go to step 4.

Step 4. Check whether x_i' maximizes $u_i(\cdot, br_j^{[0,BR_j(x_i)]}(\cdot))$ or $u_i(\cdot, br_j^{[BR_j(x_i),1]}(\cdot))$. If yes, then x is implementable. If not, then x is not implementable.

Steps 1 and 2 are easily translated into geometric analysis. An action profile can be implemented only if it lies on the best-reply curve of at least one player. If it lies on the best-reply curves of both players, this action profile is an equilibrium of the mother game, and from Proposition 1, it is implementable. Therefore, we are left with the question: *which of the action profiles that lie on only one best-reply curve can be implemented?* Steps 3 and 4 give the answer. However, these last two steps do not translate as easily into geometric analysis. In the sequel, we show that simple geometric arguments can be used to show that certain portions of the best-reply curves of the players *cannot* be implemented. Furthermore, we show that for a certain class of games, the set of implementable profiles can even be completely characterized by a straightforward geometric procedure.

To get this result, we first show that any equilibrium outcome can be described as a two step optimization program,

Proposition 5 *An outcome* (X^*, x^*) *is an equilibrium outcome of* $\Gamma(G)$ *if and only if, for at least one player* $i \in \{1, 2\}$*,* $j \neq i$*:*

(i) x_i^* *maximizes* $u_i(x_i, br_j^{X_j^*}(x_i))$ *, and*

(ii) $br_j^{X_j^*}(x_i^*) = BR_j(x_i^*)$*.*

Figure 3 illustrates the logic of Proposition 5. The outcome (x^*, X^*) with $X^* = (\{x_i^*\}, [0, \overline{x}_j])$ is an equilibrium outcome as the profile of actions x^* is associated with player i's highest indifference curves IC_i on the section of player j restricted best-reply curve $br_j^{[0,\overline{x}_j]}$ that corresponds with his mother best-reply curve BR_j. Observe that x^* is also implementable by the simple bilateral commitment $(\{x_i^*\}, [0, x_j^*])$, an illustration of Proposition 4.

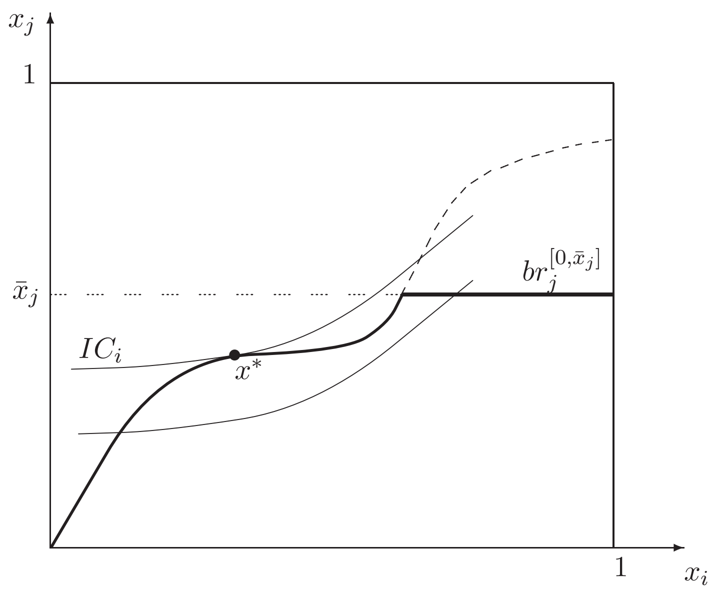

Figure 3: The geometry of Proposition 5

Remark 1 From Proposition 5, we have that x^* is implementable by the commitment X^* if x_i^* maximizes the payoff of player i being on the graph of the restricted best-reply of player j. This result has thus the flavor of the outcome of a sequential game in which player i moves first. Intuitively, this is not surprising since, as already pointed out by Schelling (1960), the power to commit oneself is equivalent to a first move.[14] Hence, implementable profiles of actions have a Stackelberg-type structure, one player 'leads' the commitment while the other 'follows.'

We now provide a geometric condition that has to hold for a profile of actions to be implementable. In other words, if this condition does not hold at a profile of actions x^* with $x_j^* = BR_j(x_i^*)$, then x^* is *not* implementable; it does not solve the above maximization program. For simplicity, assume that the (mother) best-reply maps and payoff functions are continuously differentiable.[15] The geometric condition relates the slope of the indifference curve of player i at x^* with the slope of the best-reply of player j at the same action profile x^*.

Proposition 6 *Let x^* be an implementable profile of actions with $x_j^* = BR_j(x_i^*)$, and x^* interior. It cannot be true that the slope of player i's indifference curve at x^* is strictly negative (resp., positive) while the slope of player j's (mother) best-reply at x^* is positive (resp., negative).*

Proposition 6 thus provides a general geometric condition for implementability: the slope of player i's indifference curve and the slope of player j's best-reply must have the same sign. For instance, in Figure 4, x^* is not implementable since BR_j is positively sloped at x^* while player i's indifference curve IC_i is negatively sloped. Hence, to look for implementable action profiles, we can restrict our attention to the profiles that are on the positively (resp., negatively) sloped portions of the best-reply curve of player j in the positive (resp., negative) indifference curve section of player i. This condition is not sufficient, however. In what follows, we give a necessary and sufficient geometric condition for implementation in an important class of mother games.

Consider the class of games with *strategic complementarities*.[16] Furthermore, we assume that the function $u_i(\cdot, BR_j(\cdot))$ is *strictly quasi-concave* in x_i, for all $i \in \{1, 2\}$.[17] For simplicity, we also assume that player i's payoff is increasing in player j's action x_j for all $i \in \{1, 2\}$, that

[14] There is now an abundant literature on imperfect competition whose purpose is to obtain Cournot and Stackelberg outcomes as equilibrium outcomes of the same model. Interestingly, several models use an approach similar to ours: they give the possibility to the firms to commit to some actions —see for instance Hamilton and Slutsky (1990) , van Damme and Hurkens (1999) or more recently Romano and Yildirim (2005), and the references therein. More precisely, firms in most of these models are assumed to commit either to a single action or to not commit at all. A notable exception is Romano and Yildirim (2005) who assume that firms can restrict their action sets only from the bottom, i.e., firms can only accumulate. Hence these models can be seen as a simplified version of our approach. Hamilton and Slutsky's main result is that the only equilibrium result that can be obtained are the Cournot and the Stackelberg outcomes, while our approach allows for a larger set of equilibrium results.

[15] The assumption of differentiability is not crucial, but greatly simplifies the exposition.

[16] See Fudenberg and Tirole (1991, p. 490) for a definition. It is worth noting that a similar characterization holds for games with strategic substitutabilities.

[17] See Romano and Yildirim (2005) for similar assumptions.

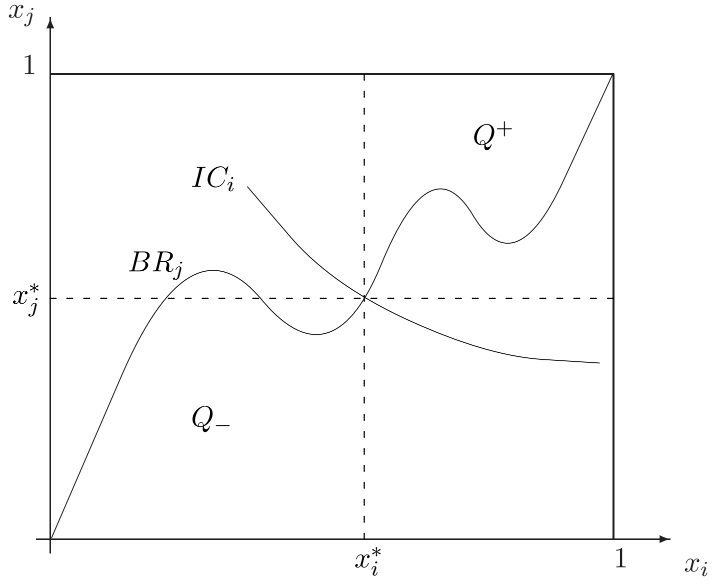

Figure 4: The profile x^* is not implementable.

is, the game has positive consonance.[18] We show that for this class of games, the knowledge of the Nash equilibria of G along with the knowledge of the 'lead-follow' profiles is necessary and sufficient to completely characterize the set of implementable profiles of actions.

First, we need to order the set of Nash equilibria of G. Define $x^*(1)$ the Nash equilibrium of G with the lowest coordinate for player i, that is, there does not exist another equilibrium x of G such that $x_i < x_i^*(1)$. Similarly, define $x^*(2)$ the equilibrium of G with the second lowest coordinate for player i, and so on recursively.[19] Note that since best-reply maps are single-valued, $x^*(k)$ is a singleton for any $k > 0$. Moreover, the set of equilibria of G is generically finite and odd (see Harsanyi (1973)), hence there generically exists a finite odd number K of $x^*(k)$'s. (See Figure 5.)

Second, define $(l_i, BR_j(l_i))$ the profile of actions such that l_i maximizes $u_i(\cdot, BR_j(\cdot))$, that is, the profile of actions $(l_i, BR_j(l_i))$ is the *lead-follow* profile with player i as the leader. It is worth noting that since $u_i(\cdot, BR_i(\cdot))$ is strictly quasi-concave in x_i and BR_j single-valued, l_i is unique. Moreover, since BR_i and u_i are non-decreasing functions of x_j, we have that $l_i \geq x_i^*(K)$ for all $i \in \{1, 2\}$ (See Lemma A3 in the Appendix). Our next proposition states that the knowledge of

[18] This assumption is not crucial. A complete characterization without this assumption is available upon request.

[19] Formally, let $x^*(0) = \emptyset$, and define for any $k > 0$,

$$x^*(k) := \{x \in N(G) \setminus \cup_{k'=0}^{k-1}\{x^*(k')\} : x_i \leq x'_i,\ \forall x' \in N(G) \setminus \cup_{k'=0}^{k-1}\{x^*(k')\}\}.$$

l_i and the $x^*(k)$'s is necessary and sufficient to completely characterize the set of implementable profiles of actions.

Before stating the proposition, let us introduce a last piece of notation. Define I_i as a subset of $[0,1]$ as follows:

$$I_i := \bigcup_{\substack{k<K \\ k \text{ odd}}} [x_i^*(k), x_i^*(k+1)] \cup [x_i^*(K), l_i]. \tag{3}$$

Observe that the set I_i is uniquely defined by the knowledge of l_i and the $x^*(k)$'s.

Proposition 7 *Consider a game with strategic complementarities and positive consonance. The set of implementable profiles of actions is $\mathcal{I} = \mathcal{I}_1 \cup \mathcal{I}_2$ with for $i \in \{1,2\}$, $j \neq i$:*

$$\mathcal{I}_i = \{x : x_j = BR_j(x_i),\ x_i \in I_i\}.$$

The intuition behind Proposition 7 is rather simple. First, note that since G is a game with strategic complementarities, the best-reply maps are increasing. Moreover, the best-reply map of any player, BR_i, separates the action space $[0,1]^2$ into two regions $\{x : x_i < BR_i(x_j)\}$ where player i's indifference curves are negatively sloped, and $\{x : x_i > BR_i(x_j)\}$ where player i's indifference curves are positively sloped. Second, for any x with $x_j = BR_j(x_i)$ and $x_i \in (x_i^*(k), x_i^*(k+1))$, k even, we have $x_i < BR_i(x_j)$, hence player i's indifference curve is negatively sloped at x. Since BR_j is positively sloped, it follows from Proposition 6 that x is *not* implementable. A similar argument holds for any x with $x_j = BR_j(x_i)$ and $x_i < x_i^*(1)$. Finally, any profile of actions x with $x_j = BR_j(x_i)$ and $x_i \in (x_i^*(k), x_i^*(k+1))$, k odd, is implementable by the simple bilateral commitment $(\{x_i\}, [0, BR_j(x_i)])$. To see this, it is enough to observe that player j's best-reply $br_j^{[0,BR_j(x_i)]}(x_i')$ is $BR_j(x_j)$ for $x_i' > x_i$, and $BR_j(x_i')$, otherwise. The strict quasi-concavity of u_i and $u_i(\cdot, BR_j(\cdot))$ implies then that x_i is solution of the optimization program described in Proposition 5. The other cases are similar. See Figure 5 for the set of implementable actions.

For the class of games with monotonic best-reply maps and $u_i(\cdot, BR_j(\cdot))$ strictly quasi-concave in x_i, the complete characterization of the set of implementable actions is therefore purely geometric, and the only knowledge required is that of the Nash equilibria of G and the lead-follow profiles.

5 The Social Value of Commitments

If we interpret our commitment game as a mechanism to implement a particular action profiles we should ask: Why don't players simply commit to efficient profile of actions? It turns out

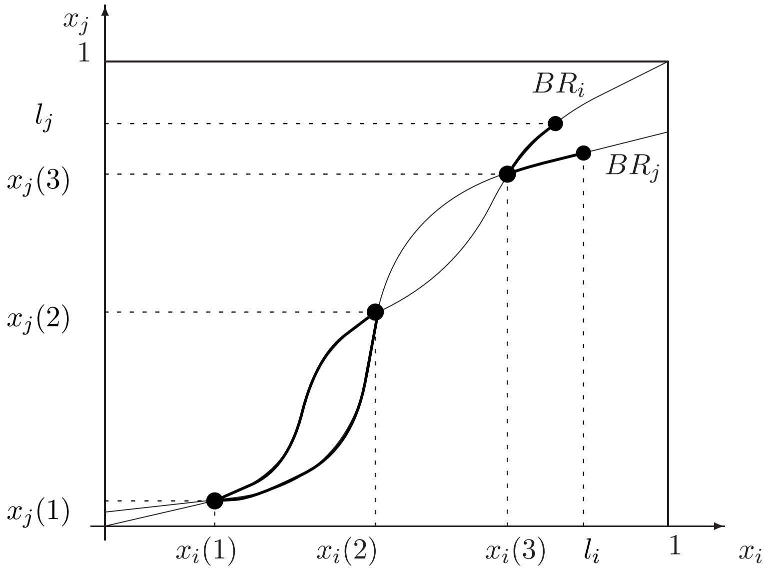

Figure 5: The set of implementable profiles (in bold)

that quite generally such commitments are not self-enforcing. More precisely, we show that if G is a smooth game, then we have generic inefficiency.

Next, we address the question of whether commitments are at least useful to implement action profiles that Pareto dominate the Nash equilibria of the mother game. We conclude, on a more positive note: we show that commitments can very well serve to make both players better off if certain conditions are met.

5.1 Efficiency

Let us first recall the definition of efficiency.

Definition 4 A profile of actions y is efficient if there does not exist another profile of actions y' such that $u_i(y') \geq u_i(y)$ for all $i \in \{1, 2\}$, and $u_i(y') > u_i(y)$ for some $i \in \{1, 2\}$.

Definition 4 is the textbook definition of (Pareto) efficiency. It is worth noting that several related papers e.g., Jackson and Wilkie (2005) or Gomez and Jehiel (2005), use a stronger concept of efficiency: a profile of actions is efficient if it maximizes the sum of players' payoffs. However, since we do not necessarily assume transferable utilities, our concept of efficiency is more appropriate. Let us now turn to the concept of smooth games.

Definition 5 The game G is a smooth game if for all $i \in N$, u_i is twice continuously differentiable.

Two remarks are in order. First, in virtually all economic models in which payoff functions are assumed to be continuous, payoff functions are also assumed to be twice continuously differentiable.[20] For instance, linear-quadratic Cournot games or models of Bertrand competition with differentiated goods are smooth games. Second, we actually need the assumption of differentiability only around equilibrium results.

Theorem 3 *For any smooth game G, interior equilibrium results of the commitment game $\Gamma(G)$ are generically inefficient.*[21]

This result is reminiscent of Theorem 1 of Dubey (1986), which states that Nash equilibria of smooth games are generically inefficient. The main reason for hope that this result could be overcome in the game of commitments is that the set of action profiles that can be implemented is (in general a large) superset of the set of Nash equilibria of the mother game. So, there is hope that this superset would also contain some efficient profiles. However, our Theorem 3 shows that this does not hold true, just like Nash equilibria of smooth games, the profiles that are implementable by commitments are generically inefficient.

Not only is our Theorem 3 reminiscent of Dubey (1986), also the proof follows along similar lines. The main difference (and difficulty) we face is that implementable profiles that are not themselves Nash equilibria of the mother game lie on the boundary of the action space of the subgame $G(X)$ with X the commitment that is implementing the profile (Lemma 2). This implies that differentiability of the restricted best response fails precisely where we need it: at the action profile under investigation.

Some additional remarks are in order. First, allowing for commitment to transfer utilities conditional on actions being played, Jackson and Wilkie (2005) also show that efficiency might not hold for two-player games. Whether efficiency holds if we allow for commitments to transfer functions and actions is an open question. Second, Theorem 3 continues to hold if G is a game with strategic complementarities, but not necessarily smooth. (See Appendix.) Third, efficient profiles on the boundary can in some games be implemented by commitments. This holds in particular if a game has an efficient Nash equilibrium on the boundary.

5.2 Pareto Improvements

While efficient results are generically not implementable, a self-enforcing commitment might nonetheless implement an improvement upon the status quo. In other words, the next question

[20]Moreover, any continuous function can be arbitrarily approximated by continuously differentiable functions by Weierstrass Approximation Theorem —See Zeidler (1986, p. 770).

[21]Let T be a set of parameters indexing the payoff functions i.e., for each player $i \in \{1, 2\}$, $u_i : X \times T \to \mathbb{R}$. By genericity, we mean that there exists an open, dense subset of T for which any equilibrium result is inefficient.

we address is whether a commitment can implement a profile that makes both players better off compared to any equilibrium of the mother game G.

Definition 6 A result x^* is an *improvement upon the status quo* if $u_i(x^*) \geq u_i(y^*)$ for all $i \in \{1, 2\}$, and $u_i(x^*) > u_i(y^*)$ for at least one player, where y^* is an action profile that is efficient in the set of mother Nash equilibria.[22]

It is not hard to find games in which improvements upon the status quo can be implemented. Just take any game with a unique Nash equilibrium y^* and a lead-follow equilibrium that dominates y^*.[23] The lead-follow equilibrium can be implemented by the commitment in which the leader restricts his action space to a singleton while the follower does not restrict his action space at all. So the more interesting question is: can commitments be used to implement improvements upon the status quo if none of the lead-follow equilibria represents such an improvement? In our next result we show that this cannot happen if the players' best responses are monotone and if the players' utilities are monotone in the actions of the opponent. We say that a game satisfies *constant consonance* if any players payoff is monotone in the action of the other player.

Theorem 4 *Let G be a game with constant consonance such that the lead-follow equilibria do not improve on the status quo. Then there exists an equilibrium improvement x^* only if at least one best-reply map is non-monotonic.*

An important implication of Theorem 4 is that if G, in addition to be a game with constant consonance is also a game with strategic complementarities or strategic substitutabilities, then commitments do only serve to improve upon the status quo if the lead-follow equilibrium is already itself such an improvement. This result sharply contrasts with Proposition 2 of Bernheim and Whinston (1989), and illustrates how seemingly innocuous restrictions on the set of feasible commitments can be critical. Bernheim and Whinston's model and our model, albeit similar in spirit, differ in two important dimensions. First, in their model only one player (the principal) has the opportunity to commit. Second, and more importantly, the principal does not only have the power to commit himself (to take a single action) but he can also restrict the action set of the other player, the agent. This contrasts with our model in which both players have the power to commit and a player can only restrict his own action set.

Theorems 3 and 4 are rather negative results in that the power of commitment does not seem to be of much social value. The following example shows that equilibrium improvements do exist

[22] Note that the set of equilibria $N(G)$ is a compact set, hence efficiency is well defined.

[23] This is the case for instance of any game with a strict second-mover advantage (e.g., differentiated Bertrand duopoly). Since the payoff of the first mover in a lead-follow profile is necessarily weakly higher than the highest Nash equilibrium, the former Pareto dominates the latter.

even in the case that neither of the lead-follow equilibria represents such an improvement.

Example 1 Take the mother game G with strategy spaces $Y_1 = Y_2 = [0, 2]$ and payoff functions:

$$u_1(y_1, y_2) = \frac{y_1}{\frac{y_1}{4} + y_2} - y_1,$$
$$u_2(y_1, y_2) = -\left(y_2 + \frac{y_1}{2} - \frac{2}{3}\right)^2 - 20y_1.$$

The best-reply map of the players are

$$BR_1(y_2) = \begin{cases} -4y_2 + 4\sqrt{y_2} & \text{if } y_2 \leq 1, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$BR_2(y_1) = \begin{cases} -\frac{1}{2}y_1 + \frac{2}{3} & \text{if } y_1 \leq \frac{4}{3}, \\ 0 & \text{otherwise.} \end{cases}$$

The mother game has a unique equilibrium, $y_1^* = 4/3(\sqrt{3}-1)$, $y_2^* = 2/3(2-\sqrt{3})$, with equilibrium payoffs of $u_i(y^*) = 4/3$, $u_j(y^*) = 80/3(1 - \sqrt{3}) \simeq -19.52$. Moreover, the lead-follow profile $(BR_1(l_2), l_2) = (1, 0)$ is associated to payoffs of $u_1((BR_1(l_2), l_2))) = 0 , u_2((BR_1(l_2), l_2)) = -1/9 \simeq -0.11$.

We now show that there exists a self-enforcing commitment which implements the action profile $\tilde{y} = (8/9, 1/9)$ with associated payoffs of $u_1(\tilde{y}) = 16/9$ and $u_2(\tilde{y}) = -1441/81 \simeq -17.79$, respectively. Clearly, both players' payoffs improve upon the Nash equilibrium. According to Proposition 2, at least one player's action must be a best-reply against the action of the other player. In the profile $\tilde{y}$, we have $8/9 = BR_1(1/9)$.

Following Proposition 4, we can focus, without loss of generality, on only two candidates for the restriction of player 1, $[0, 8/9]$ or $[8/9, 1]$. We claim that player 1's restriction cannot be $[0, 8/9]$. To see this, observe that if 1 commits to $[0, 8/9]$, then player 2 can commit to $\{1\}$ and gets a payoff of $-1/9$ (since $br_1^{[0,8/9]}(1) = 0$), which is higher than $u_2(\tilde{y})$. Therefore, the unique candidate for 1's restriction is $[8/9, 1]$. In this case, player 1's restricted best-reply is

$$br_1(y_2) = \max\{-4y_2 + 4\sqrt{y_2}, 8/9\}. \tag{4}$$

Observe that for all $y_2 \in [1/9, 4/9]$, we have $-4y_2 + 4\sqrt{y_2} \geq 8/9$. It follows that 2's payoff when $y_2 \notin [1/9, 4/9]$ is $-(y_2 - 2/9)^2 - 160/9$, which is maximized when $y_2 = 1/9$. If $y_2 \in [1/9, 4/9]$, then player 2 maximizes $u_2(y) = -4y_2 + 4\sqrt{y_2}$. That the maximum is obtained when $y_2 = 8/9$ is a simple matter of computation (albeit tedious) and is left to the reader.

6 Discussion

In this paper, we completely characterize the action profiles of strategic-form games, which are implementable by unilateral commitments. We show that an action profile is implementable *if and only if* it is implementable by *simple* commitments, i.e., commitments in which a player commits to a single action while the other player commits to a subset of actions that include his *mother* best-reply to the commitment of his opponent. In turn, this result enables us to easily characterize the implementable profiles of actions as solutions of simple optimization programs under constraints (lead-follow profiles). The complete characterization of implementable actions is our first important result. Our second important result is that the set of implementable profiles does not change if we allow the players to commit gradually. We also show that the efficient actions profiles are generically not implementable. This result sharply contrasts with the situation in which players can sign binding agreements. However, bilateral commitments can be Pareto-improving in that profiles of actions, which give a payoff higher than the best payoff in any Nash equilibrium of the mother game to each player, can be implemented. Let us now discuss some of the restrictions on the commitment technology we have considered, and how our results are likely to change with altered assumptions.[24]

6.1 Non-convex restrictions

Relaxing the assumption of convex commitments leads us to consider general commitment games.[25] To circumvent the problem of the existence of a pure Nash equilibrium in each subgame we assume that the (mother) game G is a game with strategic complementarities.[26] In this case, the set of implementable outcomes in commitment games is a subset of the set of implementable outcomes in general commitment games.[27]

Proposition 8 *Let G be a game with strategic complementarities. If x^* is implementable in a commitment game, then x^* is implementable in a general commitment game.*

We give the intuition for this result for two periods, using the case of an implementable action profile x^* with $BR_1(x_2^*) = x_1^*$. More flexibility in player 1's commitment technology does not help him: he already obtains the highest possible payoff given the commitment of player 2. As

[24]Proofs of claims made in this section are available upon request.

[25]The compactness assumption has to be retained, however. For otherwise, an equilibrium does not exist.

[26]The game G has strategic complementarities (see e.g.,Topkis (1998)) if for $y_i \geq y_i'$ and $y_j \geq y_j'$ we have $u_i(y_i, y_j) - u_i(y_i', y_j) \geq u_i(y_i, y_j') - u_i(y_i', y_j')$. Cournot duopoly, differentiated Bertrand as well as all games considered by Romano and Yildirim (2005) belong to this class of games.

[27]We say that an action profile x^* is implementable in a general commitment game if there exists a subgame-perfect equilibrium of the general commitment game with x^* being played in the last period (see Definition 2).

for player 2, the additional flexibility in his commitment technology indeed implies that he can induce a larger set of proper subgames. However, any pure Nash equilibrium of these subgames can also be obtained as an equilibrium of a game in which player 2 commits to a singleton (a convex restriction). We next identify a class of games for which allowing for non-convex commitment does not affect the set of implementable outcomes. This suggests that assuming convexity is without loss of generality in several important economic applications (including all applications considered in Romano and Yildirim).

Proposition 9 *Assume that the game G features constant externalities, the map $u_i(\cdot, BR_j(\cdot))$ is strictly quasi-concave and $T = 2$. If x^* is implementable by the general commitment X^*, then it is implementable by a simple commitment.*

6.2 Multi-dimensional action spaces

Another natural extension is to consider multi-dimensional action spaces e.g., compact-convex subsets of a n-dimensional Euclidean space. Unfortunately, we are not able to offer a definitive answer at this stage. For instance, it is not entirely clear how to translate our result about the truncation at the top or bottom to the multi-dimensional realm. What does it mean to truncate a sphere at the top? We can nonetheless offer some preliminary remarks. For example, if we assume that action spaces are Cartesian products and payoff functions strictly quasi-concave in *each* component of the multi-dimensional action of a player, all our results remain valid. In general, we need to impose stronger conditions for (most of) our characterization to hold, more particularly for Proposition 2 to remain valid. Indeed, the new complication is that a small variation in one player action, from a profile in which none of the players (mother) best reply, might now change the restricted best-reply of the other player (by moving on the boundary of its restricted action set). Additional conditions assures that this possible change does not affect too much the payoff of the deviating player.

6.3 More than two players

While a full-fledged analysis awaits future research, we can offer a preliminary observation. In the paper, we have seen that lead-follow profiles are always implementable in commitment games based on two-player nice games. We might conjecture that this result also holds for three players or more. The short answer is no. To see this, consider a Cournot triopoly game with payoff $(1 - x_i - x_j - x_k)x_i$ for firm i. If firm 1 moves first, firm 2 second and firm 3 last, the equilibrium is $(1/2, 1/4, 1/8)$. We claim that the commitment game induced by this Cournot triopoly does not have an equilibrium with outcome $(1/2, 1/4, 1/8)$. By contradiction, suppose that the triple of commitment (X_1, X_2, X_3) implements the profile $(1/2, 1/4, 1/8)$. Since the (mother) best-reply of firm 2 to $(1/2, 1/8)$ is $3/16$, we should have that the lower bound of X_2 is $1/4$ by Lemma 1. Similarly, $1/2$ has to be the lower bound of X_1. We now show that firm 1 has an incentive to deviate from X_1. If X_3 is bounded from above by $1/6$, consider a deviation by firm 1 to the commitment $\{1/2 - \varepsilon\}$. The induced game has a unique Nash equilibrium $(1/2 - \varepsilon, 1/4, 1/8)$ with a payoff to firm 1 of $(1/8 + \varepsilon)(1/2 - \varepsilon)$, a profitable deviation for ε small enough. Similarly, if X_3 is bounded from below by $a \leq 1/8$, a deviation by firm 1 to the commitment $\{1/2 + \varepsilon\}$ is profitable for sufficiently small ε. Therefore, lead-follow profiles of games with three players or more are not necessarily implementable, which suggests that the characterization of implementable profiles for such games is of a very different nature that the one we propose.

6.4 Transfers

To conclude, we mention an analogy with the literature on delegation games (see e.g., Fershtman and Judd (1987) or Kockesen and Ok (2004)). Suppose that utilities are transferable. We can then reinterpret the commitment to a set of actions as the commitment to transfer functions such that actions that a player commits not to play are strictly dominated by all actions that a player commits to play. Such transfers are equivalent to the action-forcing contracts in Fershtman and Judd (1987). Whether the assumption of transferable utilities is appropriate or not depends on the application one has in mind. A further exploration of all these themes are left for future research.

Appendix

A Characterization results

Proof of Proposition 2. The proof proceeds by contradiction. Let $s^* = (X_i^*, \sigma_i^*)_{i \in \{1,2\}}$ be an equilibrium of Γ, and suppose that (X^*, x^*) the outcome of s^* is such that $x_i^* \neq BR_i(x_j^*)$ for all $i \in \{1, 2\}$, $i \neq j$. To reach a contradiction, we first identify an action, x_1' such that $u_1(x_1', x_2^*) > u_1(x_1^*, x_2^*)$ and $br_2^{X^*}(x_1') = x_2^*$. Second, we show that there exists a strategy for player 1, s_1', such that the outcome of (s_1', s_2^*) is $(X^*, (x_1', x_2^*))$, hence a contradiction with s^* being an equilibrium.

Step 1. Since x^* is a Nash equilibrium of the game $G(X^*)$, we have $x_i^* = br_i^{X^*}(x_j^*)$ for all $i \in \{1, 2\}$, $i \neq j$. Suppose that $br_i^{X^*}(x_j^*) \neq BR_i(x_j^*)$ for all $i \in \{1, 2\}$, $i \neq j$. By continuity of BR_2 and $br_2^{X_2^*}$ (remember that $br_2^{X^*}$ is the restriction of $br_2^{X_2^*}$ to X_1^*), there exists an open interval $(x_1^* - \varepsilon, x_1^* + \varepsilon)$ with $\varepsilon > 0$ sufficiently small such that for all $x_1 \in (x_1^* - \varepsilon, x_1^* + \varepsilon)$ we have that $br_2^{X_2^*}(x_1) = x_2^*$. Next pick $\alpha \in [0, 1)$ large enough such that $x_1' = \alpha x_1^* + (1-\alpha)BR_1(x_2^*) \in (x_1^* - \varepsilon, x_1^* + \varepsilon)$. By construction of $(x_1^* - \varepsilon, x_1^* + \varepsilon)$, we have that $br_2^{X_2^*}(x_1') = x_2^*$. Moreover, $u_1(x_1', x_2^*) > u_1(x_1^*, x_2^*)$ since player 1's payoff function is strictly quasi-concave in x_1.

Step 2. We claim that the strategy $s_1' = (\{x_1'\}, \sigma_1^*)$ is a profitable deviation for player 1. The outcome of (s_1', s_2^*) is $((\{x_1'\}, X_2^*), (x_1', x_2^*))$, which, by construction, gives a strictly higher payoff to player 1. ∎

Proof of Proposition 3. Let $s^* = ((X_1^*, \sigma_1^*), (X_2^*, \sigma_2^*))$ be an equilibrium of Γ with outcome (X^*, x^*). By Proposition 2, for at least one player, say player 1, we have $x_1^* = BR_1(x_2^*)$. We claim that the strategy profile $s' := (s_1^*, s_2')$, with $s_2' = (\{x_2^*\}, \sigma_2^*)$, is also an equilibrium of Γ, with outcome $((X_1^*, \{x_2^*\}), x^*)$.

First, observe that player 1 does not have an incentive to deviate from s_1^* given player 2's strategy s_2'. Indeed, since player 2's restriction is the singleton $\{x_2^*\}$, player 1 cannot obtain a payoff higher than $u_1(BR_1(x_2^*), x_2^*)$, which is the payoff he obtains under s'. Second, to show that player 2 has no profitable deviation, we use the one shot deviation property. Since s' agrees with s^* in all proper subgames of Γ, and s^* is an equilibrium of Γ, player 2 has no profitable deviations in any of the proper subgames of Γ.

Suppose now that $s_2'' = (X_2'', \sigma_2^*)$ was a profitable deviation for player 2 given player 1' strategy s_1^*. Since player 2 is indifferent between (s_1^*, s_2') and s^*, it follows that s_2'' is also a profitable deviation from s_2^*, a contradiction with our assumption that s^* is an equilibrium. ∎

Proof of Proposition 4. Let $s^* = ((\{x_i^*\}, \sigma_i^*), (X_j^*, \sigma_j^*))$ be an equilibrium of Γ with result x^*, $X_j^* = [\underline{x}_j, \overline{x}_j]$, and $x_j^* = BR_j(x_i^*)$. Define $s_j' = ([x_j^*, 1], \sigma_j^*)$ and $s_j'' = ([0, x_j^*], \sigma_j^*)$. We claim that either (s_i^*, s_j') or (s_i^*, s_j'') is an equilibrium of Γ with result x^*. First, observe that both strategy profiles under consideration have x^* as their result. To see this, note that player i has only one action x_i^*, and player j's mother best response to x_i^*, $BR_j(x_i^*)$, is contained in his restricted action space in either case. Second, note that player j does not have an incentive to change his restricted action space given player i's commitment to $\{x_i^*\}$ as his restricted action space contains his mother best-reply $BR_j(x_i^*)$ to the single action in player 1's restricted action space .

It remains to show that player i has no profitable deviation from his commitment to $\{x_i^*\}$ given the commitment of player j to either $[x_j^*, 1]$ or $[0, x_j^*]$. Since s^* is an equilibrium of Γ, the set of action profiles that give player i a payoff strictly higher than $u_i(x^*)$, $\{x : u_i(x) > u_i(x^*)\}$, does not intersect the graph of the restricted best-reply $br_j^{[\underline{x}_j, \overline{x}_j]}$ of player j. For otherwise, player i

would have a strictly profitable deviation from s_1^*. It follows that for all $x' \in \{x : u_i(x) > u_i(x^*)\}$, we have either

$$br_j^{[\underline{x}_j, \overline{x}_j]}(x_i') - x_j' > 0, \tag{A1}$$

or

$$br_j^{[\underline{x}_j, \overline{x}_j]}(x_i') - x_j' < 0. \tag{A2}$$

We can also observe that for all $x_i \in [0, 1]$,

$$br_j^{[\underline{x}_j, \overline{x}_j]}(x_i) \leq br_j^{[x_j^*, \overline{x}_j]}(x_i) \leq br_j^{[x_j^*, 1]}(x_i),$$
$$br_j^{[\underline{x}_j, \overline{x}_j]}(x_i) \geq br_j^{[\underline{x}_j, x_j^*]}(x_i) \geq br_j^{[0, x_j^*]}(x_i).$$

Suppose that $(A1)$ holds. It follows from the above observation that for all $x' \in \{x : u_i(x) > u_i(x^*)\}$,

$$br_j^{[x_j^*, 1]}(x_i') - x_j' > 0.$$

This implies that given the commitment of player j to $[x_j^*, 1]$, player i cannot obtain a payoff strictly higher than $u(x^*)$. Therefore, player i has no profitable deviation from s_i^* given s_j', hence (s_i^*, s_j') is an equilibrium of Γ. If $(A1)$ does not hold, then $(A2)$ must hold. If $(A2)$ holds, we can use the same arguments to show that x^* is implementable by $(\{x_i^*\}, [0, x_j^*])$. ∎

Proof of Proposition 5. Observe that we can rewrite conditions *(i)* and *(ii)* as follows. A profile x^* is implementable by a bilateral commitment if and only if there exists a restriction X_j^* such that x_i^* is a solution of the following program,

$$\begin{cases} (\mathcal{P}) \begin{cases} \max_{x_i \in [0,1]} u_i(x_i, x_j) \\ \text{s.t. } x_j = br_j^{X_j^*}(x_i) \end{cases} \\ \text{such that } br_j^{X_j^*}(x_i^*) = BR_j(x_i^*), \end{cases} \tag{$\mathcal{P}^*$}$$

Note that $(\mathcal{P}^*)$ is a two-step optimization program. First, we optimize $u_i(x_i, br_j^{X_j^*}(x_i))$ with respect to x_i. This is the program $(\mathcal{P})$. Second, we check whether the solution obtained lies on the graph of j's best-reply BR_j.

($\Rightarrow$) Let $s^* = (X_i, \sigma_i^*)_{i \in \{1,2\}}$ be an equilibrium of Γ, where $X_1^* = \{x_1^*\}$. (The case when $X_2^* = \{x_2^*\}$ is symmetric). Note that we make use of Proposition 3. For all $X \in \mathcal{Y}$, the mappings σ_1^* and σ_2^* are such that $(\sigma_1^*(X), \sigma_2^*(X))$ is a Nash equilibrium of $G(X)$. In particular, if $X_1 = \{x_1\}$ for some $x_1 \in Y_1$, we have $\sigma_2^*(X) = br_2^{X_2}(x_1)$. Thus, for all deviations by player 1 to a strategy $s_1 = (\{x_1\}, \sigma_1^*)$ for some $x_1 \in Y_1$, we have $u_1(s_1, s_2^*) = u_1(x_1, br_2^{X_2^*}(x_1))$. Moreover, any deviation by player 1 to a strategy $s_1' = (X_1', \sigma_1^*)$ for some $X_1 \in \mathcal{Y}_1$ with result x is result-equivalent to a deviation of the type $s_1 = (\{x_1\}, \sigma_1^*)$ since $x_2 = br_2^{X_2^*}(x_1)$ for both profiles of strategies. Since s^* is an equilibrium, such deviations are not profitable, i.e.,

$$u_1(x_1^*, br_2^{X_2^*}(x_1^*)) \geq u_1(x_1, br_2^{X_2^*}(x_1)), \forall\, x_1 \in Y_1 \,.$$

That is, x_1^* must be a solution of $(\mathcal{P})$. By Proposition 2, we have $x_i^* = BR_i(x_j^*)$ for at least one player $i \in \{1,2\}$. Suppose that $x_2^* \neq BR_2(x_1^*)$. Then, given $(\{x_1^*\}, \sigma_1^*)$, player 2 is better-off deviating to $(\{BR_2(x_1^*)\}, \sigma_2^*)$, a contradiction with s^* being an equilibrium. Hence, we have $x_2^* = BR_2(x_1^*)$, and therefore, x_1^* is solution of $(\mathcal{P}^*)$.

($\Leftarrow$) Suppose that x_1^* is solution of $(\mathcal{P}^*)$. Consider the following strategy profile: $s_1^* = (\{x_1^*\}, \sigma_1^*)$, and $s_2^* = (X_2^*, \sigma_2^*)$, where the mappings σ_1^* and σ_2^* are such that $(\sigma_1^*(X), \sigma_2^*(X))$ is a Nash equilibrium of $G(X)$, for all $X \in \mathcal{Y}$. Clearly, the outcome of s^* is (x_1^*, x_2^*), and by construction it is a Nash equilibrium of $G(\{x_1^*\} \times X_2^*)$.[28] By construction, for all subgames $G(X)$, the actions $(\sigma_1^*(X), \sigma_2^*(X))$ constitute a Nash equilibrium of $G(X)$. Hence, according to the one-shot deviation property, it suffices to check that there is no first-stage deviation to obtain that s^* is indeed an equilibrium of Γ. Since $x_2^* = BR_2(x_1^*)$ and $X_1^* = \{x_1^*\}$, player 2 cannot obtain a better payoff than $u_2(x^*)$, and thus has no profitable deviation. As for player 1, suppose that there exists $X_1 \in \mathcal{Y}_1$ such that for $s_1 = (X_1, \sigma_1^*)$, $u_1(s_1, s_2^*) > u_1(s_1^*, s_2^*)$. Let $\tilde{x}$ be the outcome of the profile (s_1, s_2^*). Since s_1 is a profitable deviation, we then have $u_1(\tilde{x}) > u_1(x^*)$. By construction of the mapping σ_2, we have $\tilde{x}_2 = br_2^{X_2^*}(\tilde{x}_1)$, a contradiction with the fact that x_1^* is a solution of $(\mathcal{P})$. ■

B Proofs related to the multi-period game of commitments, Γ^T

Lemma A1 *Let $x^* \in N(G)$. The profile x^* is implementable in $\Gamma^T(G)$.*

Proof. The proof is similar to that of Proposition 1, and left to the reader. ■

Lemma A2 *Let x^* be implementable in $\Gamma^T(G)$. We have $x_i^* = BR_i(x_j^*)$ for at least one player $i \in \{1,2\}$.*

Proof. The proof proceeds by contradiction. Suppose that x^* is implementable in $\Gamma^T(G)$ by the strategy profile s^*, but $x_i^* \neq BR_i(x_j^*)$ for all players $i \in \{1,2\}$. Assume that $x_i^* > BR_i(x_j^*)$ for both players. (The other cases are treated similarly.) Let $s_i^*(h^t) = [\underline{x}_i^t, \overline{x}_i^t]$ where h^t is the history

[28] Since x_1^* is solution of $(\mathcal{P}^*)$, $br_2^{X_2^*}(x_1^*) = BR_2(x_1^*) \in X_2^*$. Moreover, single-valuedness of BR_2 implies that x^* is the unique Nash equilibrium of $G(\{x_1^*\} \times X_2^*)$, where $x_2^* = BR_2(x_1^*)$.

at period t on the equilibrium path. From Lemma 1 in the main text, we have that $x_i^* = \underline{x}_i^T$ for both players. Let h^{t^*} be the last history on the equilibrium path of s^* such that $\underline{x}_i^{t_i^*} \neq \underline{x}_i^T$ for at least one player $i \in \{1, 2\}$. Such an history exists as the empty history (i.e., the beginning of the game) satisfies this inequality. Without loss of generality, assume $\underline{x}_1^{t^*} \neq \underline{x}_1^T$. Moreover, as $X^t \subseteq X^{t-1}$ for any $t \in \{1, \ldots, T\}$, we have $\underline{x}_1^T > \underline{x}_1^{t^*}$, $\underline{x}_1^T = \underline{x}_1^t$ and $\underline{x}_2^T = \underline{x}_2^t$ for any $t \geq t^*+1$. We now show that player 1 has a profitable deviation at history h^{t^*}. As in the proof of Proposition 2, choose $x_1' \in (BR_1(x_2^*), x_1^*) \cap X_1^{t^*} \neq \emptyset$ sufficiently close to x_1^* such that $br_2^{X_2^{t^*+1}}(x_1') = \underline{x}_2^{t^*+1}$ where $X_2^{t^*+1}$ is the restriction played by player 2 at history h^{t^*} under s_2^*. By construction of h^{t^*}, remember that $\underline{x}_2^{t^*+1} = \underline{x}_2^T$. Construct the following strategy for player 1: $s_1'^{t^*}) = \{x_1'\}$ and $s_1'(h) = s_1^*(h)$ for any other history h. Following the history $(h^{t^*}, (\{x_1'\} \times X_2^{t^*+1}))$, the unique equilibrium result for this subgame is $(x_1', br_2^{X_2^{t^*+1}}(x_1')) = (x_1', x_2^*)$. Strict quasi-concavity of u_1 thus implies that s_1' is a profitable deviation for player 1, a contradiction. ■

Proof of Theorem 2. (⇐). The proof is trivial if $T = 1$. Suppose that $T \geq 2$. Let x^* be an action profile implementable in $\Gamma(G)$ by the simple bilateral commitment X^*. W.l.o.g. suppose that $X_1^* = \{x_1^*\}$, and $x_2^* = BR_2(x_1^*)$. We now show that we can implement x^* in $\Gamma^T(G)$. To this end, consider the strategies in $\Gamma^T(G)$ such that player 1 chooses the restriction $\{x_1^*\}$ in the first stage (and, hence in all subsequent stages) and player 2 restricts to X_2^* at the initial history and at all subsequent histories h^t of length $t < T$. Formally, we consider any profile of strategies s^* with $s_1^*(h^0) = \{x_1^*\}$ and $s_2^*(h^0) = X_2^*$ at the initial history h^0, and for any history $h^t = (h^0, (\{x_1^*\} \times X_2^*)^t)$ with $t < T$, $s_1^*(h^t) = \{x_1^*\}$ and $s_2^*(h^t) = X_2^*$. Clearly, any profile satisfying this requirement yields the result x^*. Since $x_2^* = BR_2(x_1^*)$, and given that player 1 restricts to the singleton $\{x_1^*\}$, player 2 has no incentive to deviate. As for player 1, observe that he can only deviate at the first stage. Consider a first-stage deviation by player 1 to X_1. The induced game is $\Gamma^{T-1}(G(X_1 \times X_2^*))$, and let x' be a Nash equilibrium of $G(X_1 \times X_2^*)$. By Lemma A1, there exists a profile of strategies $s^*|_{X_1 \times X_2^*}$ such that x' is implementable in $\Gamma^{T-1}(G(X_1 \times X_2^*))$, with $s^*|_{X_1 \times X_2^*}$ a profile of strategies following the history $(h^0, (X_1 \times X_2^*))$. (More precisely, let s be any profile of strategies of Γ^T, $s|_h$ is the profile of strategies induced by s after history h i.e., $s_i|_h' = s_i(h, h')$ for any h' in the set of histories following history h.) Note that since x' is the Nash equilibrium of $G(X_1 \times X_2^*)$, we have $x_2' = br_2^{X_2^*}(x_1')$, and, moreover, since $x_1^* \in \arg\max_{x_1 \in Y_1} u_1(x_1, br_2^{X_2^*}(x_1))$, we have $u_1(x^*) \geq u_1(x')$. It follows that the strategies in which player 1 commits to $\{x_1^*\}$ in the first stage, player 2 commits to X_2^* at the initial history and at all subsequent histories h^t of length $t < T$, players play $s^*|_{X_1 \times X_2^*}$ following any first-stage deviation of player 1 implements x^*. (To be complete, assume that the strategies prescribe the play of an equilibrium after any other type of histories.)

($\Rightarrow$). Let s^* be a subgame perfect equilibrium of $\Gamma^T(G)$ that implements the profile x^*, and denote (X_1^1, X_2^1) the restriction played in the first stage of $\Gamma^T(G)$. From Lemma A2, it follows that $x_i^* = BR_i(x_j^*)$ for at least one player $i \in \{1,2\}$. W.l.o.g., suppose that $x_2^* = BR_2(x_1^*)$. We claim that the commitment $(\{x_1^*\}, X_2^1)$ implements x^* in $\Gamma(G)$. Player 2 has clearly no incentive to deviate given the commitment of player 1 to $\{x_1^*\}$. Consider now player 1, and suppose that player 1 has a profitable deviation X_1' from his commitment $\{x_1^*\}$. Following player 1's deviation, the induced game is $G(X_1' \times X_2^1)$, and let x' be a Nash equilibrium of $G(X_1' \times X_2^1)$ with $u_1(x') > u_1(x^*)$. (Note that we implicitly consider the profile of strategies $((X_1', \sigma_1)(X_2^1, \sigma_2))$ with $(\sigma_1(X), \sigma_2(X))$ a Nash equilibrium of $G(X)$ for any $X \in \mathcal{Y}$.) Notice that $x_2' = br_2^{X_2^1}(x_1')$ since x' is a Nash equilibrium of $G(X_1' \times X_2^1)$. This implies that $\{x_1'\}$ is also a profitable deviation for player 1 in $\Gamma(G)$. We now show that the existence of such a deviation in $\Gamma(G)$ contradicts the fact that s^* is a subgame perfect equilibrium of $\Gamma^T(G)$. To see this, consider the strategy s_1' in which player 1 plays $\{x_1'\}$ in the first period of $\Gamma^T(G)$ and play according to s_1^* at any other history. Consider the subgame starting after this deviation by player 1. We then have the game $\Gamma^{T-1}(G(\{x_1'\} \times X_2^1))$. Clearly, in any result of this subgame player 1, plays x_1'. Therefore, the best result that player 2 can induce is $br_2^{X_2^1}(x_1')$; hence, the profile of strategies (s_1', s_2^*) leads to a unique equilibrium result, $(x_1', br_2^{X_2^1}(x_1'))$. It follows that s_1' is a profitable deviation for player 1 given the strategy s_2^* of player 2, which implies that (s_1^*, s_2^*) cannot be an equilibrium of $\Gamma^T(G)$, a contradiction. We conclude that x^* must also be implementable in $\Gamma(G)$. ∎

C Proofs related to the geometry

Proof of Proposition 6. Let x^* be an implementable profile of actions with $x_j^* = BR_j(x_i^*)$, and x^* interior. By contradiction, suppose that the slope of indifference curve of player i at x^* is negative while the slope of BR_j at x^* is positive.

Define $Q^+ := \{y \in [0,1]^2 : y \geq x^*\}$ and $Q_- = \{y \in [0,1]^2 : y \leq x^*\}$.[29] Since the indifference curve of player i at x^* is negatively sloped, there exists an $\varepsilon > 0$ such that either $u_i(y) > u_i(x^*)$ for all $y \in \mathcal{B}_\varepsilon(x^*) \cap (Q^+ \setminus \{x^*\})$ or such that $u_i(y) > u_i(x^*)$ for all $y \in \mathcal{B}_\varepsilon(x^*) \cap (Q^- \setminus \{x^*\})$, where $\mathcal{B}_\varepsilon(x^*)$ is an open ball of radius ε around x^*.

Let $f : X \to Y$ be a function. We denote $\operatorname{Gr} f$ the graph of f. Since the slope of BR_j at x^*

[29] Let x and y two vectors in $\mathbb{R}^n$. We write $x \geq y$ if $x_i \geq y_i$ for all $i \in \{1, \ldots, n\}$

is positive, we have that

$$\text{Gr } br_j^{[0,BR_j(x_i^*)]} \cap (B_\varepsilon(x^*) \cap Q^+ \setminus \{x^*\}),$$
$$\text{Gr } br_j^{[0,BR_j(x_i^*)]} \cap (B_\varepsilon(x^*) \cap Q_- \setminus \{x^*\}),$$
$$\text{Gr } br_j^{[BR_j(x_i^*),1]} \cap (B_\varepsilon(x^*) \cap Q^+ \setminus \{x^*\}),$$
$$\text{Gr } br_j^{[BR_j(x_i^*),1]} \cap (B_\varepsilon(x^*) \cap Q_- \setminus \{x^*\}),$$

are non-empty sets, hence the graph of player j's restricted best-reply intersects player i's strict upper contour set at x^*.

Finally, from Theorem 1, the two simple commitments that could possibly implement the profile x^* are $(\{x_i^*\}, [0, BR_j(x_i^*)])$ and $(\{x_i^*\}, [BR_j(x_i^*), 1])$. It follows from the above arguments that x^* cannot be a solution of the optimization program described in Proposition 5 (since the graph of player j's restricted best-reply intersects player i's strict upper contour set at x^*), hence a contradiction with x^* being implementable. The same argument follows mutatis mutandum for the other cases. ∎

Lemma A3 *Let G be a game with strategic complementarities and positive consonance i.e., u_i is non-decreasing in x_j, $j \neq i$, for all $i \in N$. We have $l_i \geq x_i^*(K)$.*

Proof. Suppose that $x_i^*(k+1) > l_i > x_i^*(k)$. Since, BR_j is non-decreasing, we have $BR_j(x_i^*(k+1)) \geq BR_j(l_i) \geq BR_j(x_i^*(k))$, hence

$$u_i(l_i, BR_j(x_i^*(k+1)) \geq u_i(l_i, BR_j(l_i)) \tag{A3}$$

since u_i has positive consonance. Moreover, since $x_i^*(k+1)$ is the unique best-reply to $x_j^*(k+1) = BR_j(x_i^*(k+1))$ ($x^*(k+1)$ is a Nash equilibrium), we have

$$\begin{aligned} u_i(x_i^*(k+1), x_j^*(k+1)) &> u_i(l_i, BR_j(x_i^*(k+1)) \\ &\geq u_i(l_i, BR_j(l_i)) \geq u_i(x_i^*(k+1), x_j^*(k+1)), \end{aligned} \tag{A4}$$

a contradiction. A similar argument shows that l_i could not be smaller than $x_i^*(1)$. ∎

Proof of Proposition 7. We first start with a preliminary observation. The best-reply of player i separates the action space $[0,1]^2$ into two regions: one region in which player i's indifference curves are negatively sloped, one region in which player i's indifference curves are positively sloped. To prove this result, fix an action x_j^* of player j, and consider the best-reply $x_i^* = BR_i(x_j^*)$ of player i to x_j^*. Define $IC := \{x \in [0,1]^2 : u_i(x) = u_i(x^*)\}$. For any $x_i \neq x_i^*$, we have $u_i(x_i, x_j^*) < u_i(x^*)$ since x_i^* is the unique best-reply to x_j^*. Next, if $x_j < x_j^*$, it follows from

u_i increasing in x_j that $u_i(x_i, x_j) \leq u_i(x_i, x_j^*) < u_i(x^*)$, hence $(x_i, x_j) \notin IC$. Therefore, for any x_i, we need $x_j > x_j^*$ for (x_i, x_j) to belong to IC. Hence, we have that for any $x_i < x_i^*$, IC is negatively sloped and for any $x_i > x_i^*$, IC is positively sloped.

As a second observation, note that for any $x_i \in [x_i^*(k), x_i^*(k+1)]$, $BR_i(BR_j(x_i)) - x_i$ is either positive or negative, but does not alternate in signs. For otherwise, there exists another equilibrium in $(x_i^*(k), x_i^*(k+1))$, a contradiction with the definition of the $x^*(k)$'s. Moreover, we have that $BR_i(BR_j(x_i)) - x_i < 0$ for any $x_i \in (x_i^*(k), x_i^*(k+1))$ if k is odd, $BR_i(BR_j(x_i)) - x_i > 0$, if k is even. In words, the graph of player i's best-reply is to the 'left' of the graph of player j's best-reply if k is odd, and to the 'right' if k is even. (See Figure 5.) Furthermore, $BR_i(BR_j(x_i)) - x_i > 0$ for any $x_i < x_i^*(1)$ and $BR_i(BR_j(x_i)) - x_i < 0$ for any $x_i > x_i^*(K)$.[30]

Fix a profile of actions x with $x_j = BR_j(x_i)$ and $x_i \in (x_i^*(k), x_i^*(k+1))$ for some k even. We want to show that this profile is not implementable. From the previous observation, we have that $BR_i(x_j) = BR_i(BR_j(x_i)) > x_i$. From the first observation, it then follows that the indifference curve of player i at x is negatively sloped. Since BR_j is positively sloped, it follows from Proposition 6 that x is not implementable. A similar argument holds for any x with $x_j = BR_j(x_i)$ and $x_i < x_i^*(1)$.

Let us now consider any profile of actions x^* with $x_j^* = BR_j(x_i^*)$ and $x_i^* \in (x_i^*(k), x_i^*(k+1))$ for some k odd. We want to show that any such a profile is implementable by the simple bilateral commitment $(\{x_i^*\}, [0, BR_j(x_i^*)])$. The key observation is that the best-reply of player i is now to the 'left' of the best-reply of player j i.e., $BR_i(BR_j(x_i^*)) < x_i^*$. (See Figure 5.) Hence, for any $x_i > x_i^*$, $br_j^{X_j^*}(x_i) = BR_j(x_i^*)$, that is, player j's restricted best-reply is $BR_j(x_i^*)$, and $u_i(x_i, br_j^{X_j^*}(x_i)) < u_i(x_i^*, br_j^{X_j^*}(x_i^*))$ by strict quasi-concavity of u_i. Finally, note that $br_j^{X_j^*}(x_i) = BR_j(x_i)$ for any $x_i \leq x_i^*$, henceforth the maximum of $u_i(\cdot, br_j^{X_j^*}(\cdot))$ is achieved in x_i^* by strict quasi-concavity of $u_i(\cdot, BR_j(\cdot))$. It follows that x^* is implementable (step 4).

Similar arguments applies to show that any point x^* with $x_j^* = BR_j(x_i^*)$ and $x_i^* \in (x_i^*(K), l_i]$ is implementable by the simple bilateral commitment $(\{x_i^*\}, [0, BR_j(x_i^*)])$. ∎

D Proofs related to the welfare

Proof of Theorem 3. Let (X^*, x^*) be any equilibrium outcome of $\Gamma(G)$ such that X^* is simple, and x^* is interior. Let T be a set of parameters and define the family of payoff functions : $u_i : X \times T \to \mathbb{R}$, for all $i \in \{1, 2\}$. We want to show that for a dense open subset T^* of T,

[30] By contradiction, suppose that $BR_i(BR_j(x_i)) - x_i < 0$ for any $x_i < x_i(1)$. In particular, for $x_i = 0$, i.e., for the lower bound of Y_i, we have $0 \leq BR_i(BR_j(0)) - 0 < 0$, a contradiction.

x^* is inefficient. If x^* is an equilibrium of the mother game G, the result follows from Theorem 1 of Dubey (1986). If x^* is not an equilibrium of the mother game G, the proof is similar to the proof of Theorem 1 of Dubey. The proof is as follows. Define the directional mapping $D : T \times X \to \mathbb{R}^4$,

$$D(t, x') = \begin{bmatrix} \frac{\partial u_1(\cdot,t)}{\partial x_1}(x') & \frac{\partial u_1(\cdot,t)}{\partial x_2}(x') \\ \frac{\partial u_2(\cdot,t)}{\partial x_1}(x') & \frac{\partial u_2(\cdot,t)}{\partial x_2}(x') \end{bmatrix}, \tag{A5}$$

and let $D_t(\cdot)$ be the restriction of D to t. Thus, $D_t(x^*)$ is the Jacobian matrix evaluated at x^*. A key step in Dubey's proof is to observe that at any interior equilibrium x^* of G, the diagonal elements of the Jacobian matrix are zero, and that the set of 2×2 matrices with zeros on the diagonal is a sub-manifold of $\mathbb{R}^4$ of co-dimension 2. If x^* is *not* an equilibrium of G, we have a similar result, that is, we can show that if x^* is an equilibrium result of Γ, then $D_t(x^*) \in A \cap B$, with $A \cap B$ a sub-manifold of $\mathbb{R}^4$ of co-dimension 2. This step is the only step that differs with Dubey's proof.

First, from Lemma 2, for at least one player, we have $x_i^* = BR_i(x_j)$. Without loss of generality, suppose that $x_2^* = BR_2(x_1^*)$. Since x^* is interior, we then have that $\frac{\partial u_2}{\partial x_2}(x^*) = 0$. This equality is our first constraint on the Jacobian matrix. Formally, define the set

$$A = \{M \in \mathbb{R}^4 : M_{22} = 0\}, \tag{A6}$$

i.e., the set of 2×2 matrices with a zero on the diagonal. Observe that if x^* is an equilibrium result, then $D_t(x^*) \in A$, or $x^* \in D_t^{-1}(A)$. The set A is a sub-manifold of $\mathbb{R}^4$ of co-dimension 1.

Second, since (X^*, x^*) is an equilibrium outcome, it follows from Theorem 1 that $u_1(x_1^*, br_2^{X_2^*}(x_1^*)) \geq u_1(x_1, br_2^{X_2^*}(x_1))$ for all $x_1 \in Y_1$. We show that these inequalities impose a relationship between the first-order derivatives of u_1 with respect to x_1 and x_2, respectively. If $br_2^{X_j^*}$ is differentiable at x^*, then the relationship is trivial. However, whenever X^* is a simple commitment, $br_2^{X_2^*}$ is not differentiable in x_1^*. We use the concepts of subgradient and subdifferential to circumvent this problem.[31]

For any function $f : Z \to \mathbb{R}$, denote $\partial f(z)$ the subdifferential of f at z. We refer the reader to Clarke (1989, Chapter 1) or Rockafellar (1981, Chapter 3) for rigorous definitions of subdifferentials. As an example, if $f(z) = |z|$, then $\partial f(0) = [-1, 1]$.

Since u_2 is twice continuously differentiable, BR_2 is continuously differentiable, hence Lipschitz continuous. From Lemma 1, it then follows that $br_2^{X_2^*}$ is Lipschitz continuous. Note that Rademacher Theorem implies that $br_2^{X_2^*}$ is differentiable almost everywhere. Let us consider the

[31] We refer the reader to Rockafellar (1981) for a good source on the theory of subgradients and non-smooth optimization.

subdiffential of $v_1(\cdot) := -u_1(\cdot, br_2^{X_2^*}(\cdot))$ at x_1^*. Since u_1 is continuously differentiable and $br_2^{X_2^*}$ is Lipschitz continuous, Theorem 5P of Rockafellar (1981, p. 74) implies that

$$\partial v_1(x_1^*) = -\frac{\partial u_1}{\partial x_1}(x^*) - \frac{\partial u_1}{\partial x_2}(x^*)\partial br_2^{X_2^*}(x_1^*). \tag{A7}$$

Since x_1^* minimizes v_1, $0 \in \partial v_1(x_1^*)$ (Clarke, 1989, p. 9)), hence there exists a $\xi \in \partial br_2^{X_2^*}(x_1^*)$ such that

$$0 = \frac{\partial u_1}{\partial x_1}(x^*) + \frac{\partial u_1}{\partial x_2}(x^*)\xi, \tag{A8}$$

the required relationship. (Note that if $br_2^{X_2^*}$ is differentiable at x_1^*, then ξ is the derivative of $br_2^{X_2^*}$ evaluated at x_1^*.)

For any scalar a, define the set

$$B = \{M \in \mathbb{R}^4 : M_{11} + aM_{12} = 0\}, \tag{A9}$$

i.e., the set of 2×2 matrices with a linear relationship between the two first entries. It follows that if x^* is an equilibrium result, then $D_t(x^*) \in B$, or $x^* \in D_t^{-1}(B)$ (take $a = \xi$). The set B is a submanifold of $\mathbb{R}^4$ of co-dimension 1. It then follows that $A \cap B$ is a submanifold of $\mathbb{R}^4$ of co-dimension 2, as required.

Finally, define the set

$$C = \{M \in \mathbb{R}^4 : \text{the rows of } M \text{ are linearly dependent}\}. \tag{A10}$$

It is easy to see that if x^* is efficient, then $D_t(x^*) \in C$, or $x^* \in D_t^{-1}(C)$. For otherwise, there exists a neighborhood O of x^* and a $x' \in O$ such that $u_i(x') = u_i(x^*) + \varepsilon_i$, $\varepsilon_i > 0$, for all player $i \in N$ i.e., there exists dx_1 and dx_2 such that

$$\begin{bmatrix} \frac{\partial u_1(\cdot,t)}{\partial x_1}(x^*) & \frac{\partial u_1(\cdot,t)}{\partial x_2}(x^*) \\ \frac{\partial u_2(\cdot,t)}{\partial x_1}(x^*) & \frac{\partial u_2(\cdot,t)}{\partial x_2}(x^*) \end{bmatrix} \begin{pmatrix} dx_1 \\ dx_2 \end{pmatrix} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}. \tag{A11}$$

Hence, if a profile x^* is an equilibrium result and efficient, then $D_t(x^*) \in A \cap B \cap C$ or $x^* \in D_t^{-1}(A \cap B \cap C)$.

The next step is to show that for a dense open set $T^* \subset T$, $D_t^{-1}(A \cap B \cap C)$ is empty. To do so, we shall show that the co-dimension of $D_t^{-1}(A \cap B \cap C)$ is 2, that is the dimension of Y, hence is empty. This step is found in Dubey's proof. ∎

Inefficiency and a non-smooth game

Assume that the game G is a game with strategic complementarities and negative consonance i.e., $x_j \mapsto u_i(x_i, x_j)$ is decreasing in x_j for each player $i \in N$, $i \neq j$. Note that G is not assumed to be smooth.

The first observation is that $BR_1(BR_2(x_1^*)) \leq x_1^*$. Since BR_2 is monotone increasing in x_1, we have $br_2^{[0,BR_2(x_1^*)]}(x_1) = BR_2(x_2)$ for all $x_2 \in [0, x_1^*]$, and $br_2^{[0,BR_2(x_1^*)]}(x_1) = BR_2(x_1^*)$, otherwise. Henceforth, if $BR_1(BR_2(x_1^*)) > x_1^*$, we have that player 2's best-reply to $BR_1(BR_2(x_1^*))$ is $BR_2(x_1^*)$, hence a contradiction with x_1^* maximizing player 1's payoff on the constrained best-reply of player 2.

Second, since u_2 is decreasing in x_1, we obviously have

$$u_2(x_2^*, BR_1(BR_2(x_1^*))) \geq u_2(x_2^*, x_1^*),$$

hence $(BR_1(BR_2(x_1^*)), x_2^*)$ improves upon 2's payoff.

Finally, since at an equilibrium x^* of Γ, $x_2^* = BR_2(x_1^*)$, it follows that

$$u_1(BR_1(x_2^*), x_2^*) \geq u_1(x^*),$$

with a strict inequality if x^* is not a Nash equilibrium of G.

It follows that $(BR_1(x_2^*), x_2^*)$ Pareto-improves upon x^*, hence x^* is not efficient. Finally, observe that the result also holds if we assume strategic substitutes and payoff increasing in the action of the opponent.

Proof of Theorem 4 Let (X^*, x^*) be an equilibrium outcome of Γ and assume that x^* is an improvement upon the status quo. Let x^N be a Nash equilibrium, which is efficient in the set of Nash equilibria, for which we have $u_i(x^*) \geq u_i(x^N)$ for $i \in \{1, 2\}$ with at least one strict inequality. Using Proposition 2, we can assume that $x_2^* = BR_2(x_1^*)$. By our assumption that neither of the lead-follow equilibria is an improvement upon the status quo, we have that

$$u_2(x^*) \geq u_2(x^N) > u_2(l_1, BR_2(l_1)).$$

Observe that in all the three profiles, player 2 is best replying to player 1's action. Furthermore, as player 2's payoff function is monotonic in his opponent's action, we have that $u_2^*(x_1) := u_2(x_1, BR_2(x_1))$ is a monotonic function of x_1, hence x_1^* and l_1 must lie on two different sides of x_1^N i.e., we must have either $l_1 \geq x_1^N \geq x_1^*$ or $l_1 \leq x_1^N \leq x_1^*$. Since best-reply maps are single valued, we also have that $l_1 \neq x_1^N \neq x_1^*$.

Moreover, since x^N and $(l_1, BR_2(l_1))$ both lie on the graph of player 2's mother best-reply and u_1 is continuous, we have

$$u_1(l_1, BR_2(l_1)) \geq u_1(x^*) \geq u_1(x^N)\,.$$

Assume that player 2's best-reply function is monotonic. We will show that l_1 and x_1^* cannot lie on two different sides of x_1^N, and give to player 1 a payoff higher than his Nash payoff whenever player 1's payoff function is monotonic in his opponent's action and best-reply functions are monotonic.

We first start with the case in which the best-reply function BR_2 is non-decreasing and the player 1's payoff function has positive consonance i.e., $x_2 \mapsto u_1(x_1, x_2)$ is non-decreasing. From Lemma A3, we have $l_1 > x_1^N$, therefore $l_1 > x_1^N > x_1^*$ since l_1 and x_1^* must lie on two different sides of x_1^N. Moreover, $BR_2(x_1^N) \geq BR_2(x_1^*)$. It thus follows that

$$u_1(x_1^N, BR_2(x_1^N)) > u_1(x_1^*, BR_2(x_1^N)) \geq u_1(x_1^*, BR_2(x_1^*)),$$

where the first strict inequality follows by strict quasi-concavity and the second by positive consonance, a contradiction.

Second, consider the case in which the best-reply function BR_2 is non-decreasing and the player 1's payoff function has negative consonance i.e., $x_2 \mapsto u_1(x_1, x_2)$ is non-increasing. An immediate modification of Lemma A3 implies that $l_1 < x_1^N$, and therefore $l_1 < x_1^N < x_1^*$. It follows that $BR_2(x_1^N) \leq BR_2(x_1^*)$, and

$$u_1(x_1^N, BR_2(x_1^N)) > u_1(x_1^*, BR_2(x_1^N)) \geq u_1(x_1^*, BR_2(x_1^*)),$$

where the first strict inequality follows by strict quasi-concavity and the second by negative consonance, a contradiction.

The other cases are similar and left to the reader. ∎

E Proofs related to general (non-convex) commitments

Proof of Proposition 8. Let $T = 2$. Assume that the profile x^* is implementable by X^*. Since x^* is implementable, it follows from Theorem 1 that x^* is implementable by the commitment $(\{x_i^*\}, X_j^*)$ with $x_j^* = BR_j(x_i^*)$. Let us show that x^* is also implementable in a general commitment game. First, since player j's payoff is the highest payoff player j can obtain when player i is committed to x_i^*, player j has no profitable deviation. Second, if player i deviates to any general commitment X_i (non-necessarily convex), the induced game is $G(X_i \times X_j^*)$. Since this game is a game with strategic complementarities, it has a pure Nash equilibrium $\tilde{x} := (\tilde{x}_i, br_j^{X_j^*}(\tilde{x}_i))$ with a payoff to player i of

$$u_i(\tilde{x}) = u_i(\tilde{x}_i, br_j^{X_j^*}(\tilde{x}_i)) \leq \max_{x_i \in [0,1]} u_i(x_i, br_j^{X_j^*}(x_i)) = u_i(x^*),$$

hence the deviation is not profitable.

Let $T > 2$. To implement x^* in the general commitment game, consider the following strategies. Player i commits to $\{x_i^*\}$ and player j to X_j^* in the first period. In period $t > 1$, player j continue to play X_j^* if he has always played X_j^* in the previous $t-1$ periods. Player i has no choice but to play $\{x_i^*\}$. Following a deviation to X_j^t by player j at period t, player j plays $\{\tilde{x}_j\}$ where $(x_i^*, \tilde{x}_j)$ is a pure Nash equilibrium of $G(\{x_i^*\} \times X_j^t)$. If either player deviates in the first period, the strategies require them to commit to a pure Nash equilibrium of the game induced by their deviation in the next period. In the last period, the strategies prescribe the play of a pure Nash equilibrium. It follows from the proof of Proposition 1 and the above arguments that x^* is implementable. ∎

Proof of Proposition 9. Assume that the game features positive externalities (the arguments are similar if we assume negative externalities). Let x^* be implementable by the general commitment (X_i^*, X_j^*) and assume that $x_j^* < BR_j(x_i^*)$. Let $X_j^{**} := [\min X_j^*, x_j^*]$, it is convex. We first show that x^* is implementable by the commitment (X_i^*, X_j^{**}). By strict quasi-concavity, we have that x^* is a Nash equilibrium of $G(X_i^* \times X_j^{**})$ (See Lemma 1). Moreover, since x^* is implementable by X^*, player j has no incentive to deviate from X_j^{**}. To see this, suppose that player j has a profitable deviation i.e., there exists a commitment X_j such that *all* equilibria of $G(X_i^* \times X_j)$ gives player j a payoff strictly higher than $u_j(x^*)$. Then, we have a contradiction with x^* being implementable by X^* since subgame perfection requires to play a Nash equilibrium of $G(X_i^* \times X_j)$. All these equilibria would give to player j a payoff strictly higher than $u_j(x^*)$. Similarly, the graph of $br_j^{X_j^*}$ is included in the lower contour set $LC_i(x^*)$ of player i at x^* with $LC_i(x^*) := \{x \in [0,1]^2 : u_i(x) \leq u_i(x^*)\}$. It remains to show that the graph of $br_j^{X_j^{**}}$ is also included in $LC_i(x^*)$. Loosely speaking, we want to show that player i cannot "move" along the graph of player j's restricted best-reply and find a profile that strictly improves his payoff over $u_i(x^*)$. For otherwise, he would have a profitable deviation.

First, for all $x_i \in [0,1]$ such that $br_j^{X_j^{**}}(x_i) = br_j^{X_j^*}(x_i)$, we clearly have $(x_i, br_j^{X_j^{**}}(x_i)) \in LC_i(x^*)$. Second, for all $x_i \in [0,1]$ such that $br_j^{X_j^{**}}(x_i) < br_j^{X_j^*}(x_i)$, we have that

$$u_i(x_i, br_j^{X_j^{**}}(x_i)) \leq u_i(x_i, br_j^{X_j^*}(x_i)) \leq u_i(x^*),$$

where the first inequality comes from positive externalities and the second from the fact that the graph of $br_j^{X_j^*}$ is in $LC_i(x^*)$. Third, consider all $x_i \in [0,1]$ such that $br_j^{X_j^{**}}(x_i) > br_j^{X_j^*}(x_i)$. We have that

$$u_i(x_i, br_j^{X_j^*}(x_i)) \leq u_i(x_i, br_j^{X_j^{**}}(x_i)) \leq u_i(x_i, x_j^*),$$

where the last equality comes from $x_j^* \geq br_j^{X_j^{**}}(x_i)$ for all $x_i \in [0,1]$. Suppose that there exists a $\tilde{x}_i$ with $br_j^{X_j^{**}}(\tilde{x}_i) > br_j^{X_j^*}(\tilde{x}_i)$ such that $u_i(\tilde{x}_i, x_j^*) > u_i(x_i^*, x_j^*)$. Since BR_j is increasing and $x_j^* < BR_j(x_i^*)$, there either exists a $\hat{x}_i < x_i^*$ such that $BR_j(\hat{x}_i) = x_j^*$ or $BR_j(x_i) > x_j^*$ for all $x_i \in [0,1]$. Consider the former case. We have that $\tilde{x}_i < \hat{x}_i$. To see this, note that $br_j^{X_j^{**}}(x_i) = x_j^*$

for all $x_i \geq \hat{x}_i$ since best-replies are increasing. Moreover, assume that there exists a $\tilde{x}_i \geq \hat{x}_i$ such that $br_j^{X_j^{**}}(\tilde{x}_i) > br_j^{X_j^*}(\tilde{x}_i)$. This implies that $BR_j(\tilde{x}_i) \geq BR_j(\hat{x}_i) = x_j^* > br_j^{X_j^*}$, a contradiction with the strict-quasi concavity of the payoff function and $x_j^* \in X_j^*$. Consequently, the strict quasi-concavity of u_i implies that $u_i(x_i, x_j^*) > u_i(x^*)$ for all $x_i \in [\hat{x}_i, x_j^*]$, a contradiction since these points belong to the graph of $br^{X_j^*}$. Consider now the latter case i.e., $BR_j(x_i) > x_j^*$ for all $x_i \in [0,1]$. In this case, we show that $br_j^{X_j^{**}}(x_i) \leq br_j^{X_j^*}(x_i)$ for all $x_i \in [0,1]$. To see this, suppose there exists a $\hat{x}_i$ such that $br_j^{X_j^{**}}(\hat{x}_i) > br_j^{X_j^*}(\hat{x}_i)$. However, strict quasi-concavity implies that $br_j^{X_j^{**}}(x_i) = x_j^*$ for all x_i, henceforth $BR_j(\hat{x}_i) > br_j^{X_j^{**}}(\hat{x}_i) = x_j^* > br_j^{X_j^*}(\hat{x}_i)$, a contradiction since $x_j^* \in X_j^*$ and u_i is strictly quasi-concave in x_i. Henceforth, x^* is implementable by the commitment (X_i^*, X_j^{**}).

If $x_j^* > BR_j(x_i^*)$, apply the same arguments as above with $X_j^{**} := [x_j^*, \max X_j^*]$.

If $x_j^* = BR_j(x_i^*)$, strict quasi-concavity of the map $x_i \mapsto u_i(x_i, BR_j(x_i))$ implies that x^* is implementable by either $(\{x_i^*\}, [0, BR_j(x_i^*)])$ or $(\{x_i^*\}, [BR_j(x_i^*), 1])$, two simple commitments. For instance, suppose that the lead-follow profile $(l_i, BR_j(l_i))$ is higher than $(x_i^*, BR_j(x_i^*))$, then x^* is implementable by the simple commitment $(\{x_i^*\}, [0, BR_j(x_i^*)])$. To see this, observe that $br_j^{[0,BR_j(x_i^*)]}(x_i) = BR_j(x_i)$ for any $x_i \leq x_i^*$ and $br_j^{[0,BR_j(x_i^*)]}(x_i) = x_j^*$ for any $x_i > x_i^*$ since BR_j is increasing. Clearly, strict quasi-concavity of the map $x_i \mapsto u_i(x_i, BR_j(x_i))$ implies that player i has no profitable deviation to $\tilde{x}_i \leq x_i^*$. Suppose he has a profitable deviation to $\tilde{x}_i > x_i^*$. First, if $(\tilde{x}_i, x_j^*)$ belongs to the graph of $br_j^{X_j^*}$, we have a contradiction. Second, if $(\tilde{x}_i, x_j^*)$ does not belong to the graph of $br_j^{X_j^*}$, then $br_j^{X_j^*}(\tilde{x}_i) > x_j^*$. It follows from positive externalities that

$$u_i(\tilde{x}_i, br_j^{X_j^*}(\tilde{x}_i)) \geq u_i(\tilde{x}_i, x_j^*) > u_i(x^*),$$

again a contradiction with the implementation of x^* by the general commitment X^*. To complete the proof, repeat the same arguments as above but starting with x^* being implementable by the commitment (X_i^*, X_j^{**}). ∎

References

[1] A. R. Admati and M. Perry, *Joint projects without commitment*, Review of Economic Studies, 58 (1991), pp. 259–276.

[2] R. Amir and I. Grilo, *Stackelberg versus Cournot equilibrium*, Games and Economic Behavior, 26 (1999), pp. 1–21.

[3] S. Bade, G. Haeringer, and L. Renou, *More strategies, more Nash equilibria*. Journal of Economic Theory. Forthcoming. August 2005.

[4] J. Brander and B. Spencer, *Export subsidies and international market share rivalry*, Journal of International Economics, 18 (1985), pp. 83–100.

[5] B. D. Bernheim and M. D. Whinston, *Incomplete contracts and strategic ambiguity*, American Economic Review, 88 (1998), pp. 902–932.

[6] G. Caruana and L. Einav, *A theory of endogenous commitment*, Mimeo, June 2005.

[7] F. H. Clarke, *Methods of dynamic and nonsmooth optimization*, CBMS-NSF Regional Conference Series in Applied Mathematics; 57, Society for Industrial and Applied Mathematics, 1989.

[8] E. van Damme and S. Hurkens, *Endogenous Stackelberg leadership*, Games and Economic Behavior, 28 (1999), pp. 105–129.

[9] A. Dixit, *The role of investment in entry-deterrence*, Economic Journal, 90 (1980), pp. 95–106.

[10] P. Dubey, *Inefficiency of Nash equilibria*, Mathematics of Operations Research, 11 (1986), pp. 1–8.

[11] C. Fershtman and K. Judd, *Equilibrium incentives in oligopoly*, American Economic Review, 77 (1987), pp. 927–940.

[12] D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, Cambridge, 1991.

[13] D. Gale, *Monotone games with positive spillovers*, Games and Economic Behavior, 37 (2001), pp. 295–320.

[14] A. Gomez and P. Jehiel, *Dynamic processes of social and economic interactions*, Journal of Political Economy, 113 (2005), pp. 626–667.

[15] J. H. Hamilton and S. M. Slutsky, *Endogeneous timing in duopoly games: Stackelberg or Cournot equilibria*, Games and Economic Behavior, 2 (1990), pp. 29–46.

[16] C. Harris, P. J. Reny, and A. Robson, *The existence of subgame-perfect equilibrium in continuous games with almost perfect information: A case for extensive-form correlation*, Econometrica, 63 (1995), pp. 507–544.

[17] J. Harsanyi, *Oddness of the number of equilibrium points: a new proof*, International Journal of Game Theory, 2 (1973), pp. 235–250.

[18] J. Henkel, *The 1.5th mover advantage*, RAND Journal of Economics, 33 (2002), pp. 156–170.

[19] O. Hart and J. Moore, *Agreeing now to agree later: Contracts that rule out but do not rule in.* Mimeo, London School of Economics, May 2004.

[20] M. Jackson and S. Wilkie, *Endogenous games and mechanisms: Side payments among players*, Review of Economic Studies, 72 (2005), pp. 543–566.

[21] L. Kockesen and E. Ok, *Strategic delegation with unobservable incentive contracts*, Review of Economic Studies, 71 (2004), pp 118–133.

[22] D. Kreps and J. Scheinkman *Quantity precommitment and Bertrand competition yields Cournot outcomes*, Bell Journal of Economics, 14 (1983), pp. 326–337.

[23] B. Lockwood and J. P. Thomas, *Gradualism and irreversibility*, Review of Economic Studies, 69 (2002), pp. 339–356.

[24] H. Moulin, *Dominance solvability and Cournot stability*, Mathematical Social Sciences, 7 (1984), pp. 83–102.

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

[26] L. Renou, *Commitment Games*, Mimeo, University of Leicester, 2006.

[27] R. Rockafellar, *The Theory of subgradients and its applications to problems of optimization. Convex and nonconvex funtions*, Research and Education in Mathematics, Heldermann Verlag Berlin, 1981.

[28] R. Romano and H. Yildirim, *On the endogeneity of Cournot-Nash and Stackelberg equilibria: games of accumulation*, Journal of Economic Theory, 120 (2005), pp. 73–107.

[29] G. Saloner, *Cournot duopoly with two production periods*, Journal of Economic Theory, 42 (1987), pp. 183–187.

[30] D.F. Spulber *Capacity, output and sequential entry*, American Economic Review, 71 (1981), pp. 503–514.

[31] T. C. Schelling, *An essay on bargaining*, American Economic Review, 46 (1956), pp. 281–306.

[32] ——, *The Strategy of Conflict*, Harvard University Press, Cambridge, MA, 1960.

[33] E. Zeidler, *Nonliner functional analysis and its applications I: Fixed-Point Theorems*, Springer-Verlag, New-York, 1986.

[34] H. Yildirim *Contests with multiple rounds*, Games and Economic Behavior, 51 (2005), pp. 213–227.