Coalition-Stable Equilibria in Repeated Games

Anthony Fai-Tong Chung
Department of Economics
Stanford University *

This version: January, 2004

Do Not Circulate Without Permission
Comments Are Welcome

Abstract

It is well-known that subgame-perfect Nash equilibrium does not eliminate incentives for joint-deviations or renegotiations. This paper presents a systematic framework for studying non-cooperative games with group incentives, and offers a notion of equilibrium that refines the Nash theory in a natural way and answers to most questions raised in the renegotiation-proof and coalition-proof literature. Intuitively, I require that an equilibrium should not prescribe in any subgame a course of action that some coalition of players would jointly wish to deviate, given the restriction that every deviation must itself be self-enforcing and hence invulnerable to further self-enforcing deviations.

The main result of this paper is that much of the strategic complexity introduced by joint-deviations and renegotiations is redundant, and in infinitely-repeated games with discounting every equilibrium outcome can be supported by a stationary set of optimal penal codes as in Abreu (1988). In addition, I prove existence of equilibrium both in stage games and in repeated games, and provide an iterative procedure for computing the unique equilibrium-payoff set.

*Corresponding e-mail: faichung@stanford.edu. The author thanks Kenneth Arrow, Douglas Bernheim, Peter DeMarzo, Peter Hammond, Kenneth Judd, Jonathan Levin, Thomas Sargent, Ilya Segal, Mark Wright and Winnie Choi for many valuable comments. The author would especially like to thank Robert Hall for ongoing support, Paul Milgrom and Narayana Kocherlakota for many insightful and detailed discussion of various versions of this paper. The author gracefully acknowledges financial support from Stanford Economics Department and the Hoover Institution.

1 Introduction

Over the past two decades, many papers have attempted to define and characterize the set of self-enforcing agreements among rational individuals that could withstand the possibility of *renegotiation*: that after some history participants may jointly prefer to scrap the original implicit contract and set up a new arrangement. Among the first are papers by Farrell and Maskin (1989) and Bernheim and Ray (1989), who pointed out that the common practice of selecting the Pareto-optimal perfect Nash equilibria lacks "internal consistency". They questioned that if individuals can coordinate on an efficient equilibrium at the beginning of time, why would they submit to an inefficient equilibrium in some subgames, when an alternative Pareto-improving equilibrium is available? Their criticism is especially severe in situations where optimal equilibria are supported by punishments that hurt both the innocent and the guilty.

Since then, several notions of "renegotiation-proof" equilibrium have emerged from the literature.[1] Nevertheless, progress in understanding these equilibria has been impeded by the fact that, until recently, it is still unclear what constitutes a "credible" group-deviation or renegotiation. Moreover, many papers in the literature had adopted a non-behavioral approach. Intuitive properties of "renegotiation-proof" equilibria are first identified and taken as primitives in the construction of the corresponding solution concept. This obscures the link between the solution concept and the corresponding restrictions imposed on beliefs and behavior of rational individuals, and makes comparison between different solution concepts difficult. Finally, existing solution concepts focus mostly on renegotiations initiated by the grand-coalition. The possibility that members of a sub-coalition may renegotiate among themselves further complicates equilibrium characterization, and undermines the applicability of these concepts in games with more than two players.[2]

This paper presents a systematic framework for studying non-cooperative games with group incentives, and extends the notion of *coalition-stable equilibria* presented in Chung (2004) to repeated games, where I showed applying forward-induction logic in pre-play communication stage allows players to correlate their strategies, and exercise a form of coalitional reasoning. The theory developed here is also related on the seminal work of Bernheim et al. (1987), who argued that the only credible threats to a self-enforcing agreement are deviations that are themselves self-enforcing. Intuitively, I require that an equilibrium should not prescribe in any subgame a course of action that some coalition of players would jointly wish to deviate, given the restriction that every deviation must itself be self-enforcing and hence invulnerable to further self-enforcing deviations.

Despite their similarity in intuitive characterization, coalition-stable equilibria is very different

[1] See for example Farrell and Maskin (1989), Bernheim and Ray (1989), Pearce (1987), Asheim (1991), DeMarzo (1992), Bergin and MacLeod (1993), Abreu, Pearce and Stacchetti (1993), Ray (1994) and many others.

[2] Few exceptions include Bernheim, Peleg and Whinston (1987) and DeMarzo (1992) who defined solution concepts that account for deviations by sub-coalitions.

from the solution concept proposed by Bernheim et al. (1987), and is capable of resolving several conceptual difficulties left unanswered in their work. First, since my theory is motivated by a model of public pre-play communication, it removes the restriction in Bernheim et al. (1987) that only members of the deviating coalition could contemplate further deviations.[3,4] Second, because players coordinate their equilibrium behavior as well as any coalitional deviation from equilibrium using the same pre-play communication mechanism, I am able to prove generic existence of coalition-stable equilibria in mixed strategies. Finally, as I will focus in this paper, the concept of coalition-stable equilibria can be easily extended to both finitely and infinitely repeated games.

The main result of this paper is that much of the strategic complexity introduced by joint-deviations and renegotiations is redundant, and in infinitely-repeated games with discounting, every equilibrium outcome can be supported by a stationary set of optimal penal codes as in Abreu (1988). In particular, I extend the well-known "no-gain-from-one-shot-deviation" principle of Abreu (1988) and Harris (1985) to the domain of self-enforcing coalitional deviations. Hence to prevent joint-deviations or renegotiations from equilibrium, it suffices to punish one of the participants with her worst coalition-stable equilibrium, immediately after the first period of deviation.[5] The incentive structure of coalition-stable equilibrium guarantees both the punishers and the punished prefer to follow equilibrium recommendation after every history. In section 6, I'll demonstrate how to support collusion in an infinitely-repeated model of Cournot duopoly with renegotiation by the use of optimal penal codes.

Finally, for infinitely-repeated games, this paper generalizes the iterative procedure developed by Abreu, Pearce and Stacchetti (1990) to calculate the payoff set associated with coalition-stable equilibria. Let V^p be the payoff set associated with strategy profiles that is immune to self-enforcing deviations by dynamic coalitions of size less than or equal to p.[6] Through a succession of propositions, I show that for each $p = 1, ..., n$, there exists a monotone set-valued operator $B_p(\cdot)$ such that, given V^{p-1}, we can successively approximate V^p by iterating over B_p until convergence, i.e. $V^p = \lim_{k\to\infty} B_p^k(V^{p-1})$.[7] Besides offering computational tractability, this iterative procedure also leads to an existence proof of equilibrium in infinitely-repeated games.

The paper is organized as follows. Section 2 covers basic setup and notations. Section 3 reviews the development of the equilibrium concept in stage games. Section 4 motivates and extends the

[3]Bernheim et al. (1987) justified this restriction on acceptable further deviations by assuming that members of a coalition can *secretly* coordinate their moves prior to the play of the game. More importantly, members of a coalition can *commit* not to reveal their coalitional agreement to non-members.

[4]Also see Kaplan (1992) and Milgrom and Roberts (1996) for a related solution concept that partially relaxes the restriction on acceptable further deviations imposed by Bernheim et al. (1987).

[5]DeMarzo (1992) called a strategy of this form a "scapegoat" strategy, which he used to characterize his concept of *sustainable social norm* for finitely repeated games.

[6]I use V^0 to denote the payoff profiles associated with all feasible strategies, and V^n is the payoff set for coalition-stable equilibria in an n-player game.

[7]When $p = 1$, $B_1(\cdot)$ is identical to the $B(\cdot)$ operator in Abreu et al. (1990).

solution concept to repeated games. Section 5 provides a recursive characterization of equilibrium in infinitely repeated games. Section 6 applies the equilibrium concept and solves for a simple repeated Cournot duopoly model with renegotiation. Section 7 relates the theory with other solution concepts in the literature and concludes. All proofs are presented in the Appendix.

2 Setup and Notations

2.1 The Stage Game

The stage game is denoted by $G = (\{A_i\}_{i=1}^n; \{u_i\}_{i=1}^n)$, where $\mathcal{N} = \{1, ..., n\}$ is the set of players, A_i is the strategy set for player i, and $u_i : A_1 \times A_2 \times ... \times A_n \to R$ is player i's payoff function. I assume that A_i is a nonempty compact Euclidean space, and u_i is continuous and bounded for each i.[8] Elements of A_i are denoted a_i and are referred to as *actions*. Finally, the space of feasible strategy profiles is denoted by $\Sigma^0 = A \equiv A_1 \times A_2 \times ... \times A_n$, and I set $a \equiv (a_1 \times a_2 \times ... \times a_n)$, and $u(a) \equiv (u_1(a), u_2(a), ..., u_n(a))$.

2.2 The Repeated Game

Time starts at $t = 0$. Let $G^T(\delta)$ denote the repeated game generated by repeating G for time periods $0, ..., T$, where T is the horizon which can either be finite or infinite. I assume players evaluate payoffs using a common discount factor $\delta \in (0, 1)$. More precisely, for any *action profile* $\alpha \equiv \{a_t\}_{t=0}^T$ where $a_t \in A$, player i's average discounted payoff is given by $U_i(\alpha) = \frac{(1-\delta)}{(1-\delta^{T+1})} \sum_{t=0}^T \delta^t u_i(a_t)$, and is equal to $(1 - \delta) \sum_{t=0}^\infty \delta^t u_i(a_t)$ when $T = \infty$.[9]

Let h_t denote a *t-history* of all players' actions up to but not including period t, and let H_t be the set of all t-histories. A strategy for player i, denoted by σ_i, is a sequence of functions $\{\sigma_{i,t}\}_{t=0}^T$ such that $\sigma_{i,0} \in A_i$ and $\sigma_{i,t} : H_t \to A_i$ $\forall t \geq 1$.[10] I write the restriction of σ to h_t as $\sigma|_{h_t}$. The set of player i's feasible strategies is again denoted by Σ_i^0, and $\Sigma^0 \equiv \Sigma_1^0 \times \Sigma_2^0 \times ... \times \Sigma_n^0$ is the set of all feasible strategy profiles.

Given $\sigma \equiv \{\sigma_i\}_{i\in\mathcal{N}} \in \Sigma^0$, let $\alpha(\sigma)$ denote the action profile induced by σ.[11] More generally,

[8] More generally, I only need to assume that A_i is a nonempty compact Hausdorff topological space for each i. Hence the results developed here apply to all finite-action normal-form stage games.

[9] Since u_i is continuous and bounded, $\delta \in (0, 1)$ and A is compact, $U_i(\cdot)$ is uniformly continuous.

[10] When $\sigma = \{\sigma_i\}_{i\in\mathcal{N}}$ is interpreted as a behavior strategy profile, the notation employed here implies players can observe *ex-post* the realization of every player's private randomization device. Nevertheless, the analysis could be easily extended to the case where histories consist only of all previous actions taken by players, if attention is restricted to public perfect equilibria.

[11] Given $\sigma \in \Sigma^0$, the action profile $\{a_t\}_{t=0}^T = \alpha(\sigma)$ can be constructed by setting $a_0 = \{\sigma_{i,0}\}_{i\in\mathcal{N}}$ and $a_t = \{\sigma_{i,t}(h_t)\}_{i\in\mathcal{N}}$ $\forall t \geq 1$ where $h_t = (a_0, ..., a_{t-1})$.

given any strategy profile $\sigma \in \Sigma^0$ and t-history h_t, I define $\alpha(\sigma, h_t)$ to be the action profile that is equal to h_t up to time $t-1$, and is determined by subsequent applications of σ thereafter. Let $v_i(\sigma) = U_i(\alpha(\sigma))$ be player i's value function associated with σ, and let $v_i(\sigma, h_t) = U_i(\alpha(\sigma, h_t))$.[12] I use $v(\sigma)$ and $v(\sigma, h_t)$ to denote $\{v_i(\sigma)\}_{i\in\mathcal{N}}$ and $\{v_i(\sigma, h_t)\}_{i\in\mathcal{N}}$, respectively.

Throughout, I will use the notation that for any n-tuple $(x_1, ..., x_n)$, $x_s \equiv \{x_i\}_{i\in s}$ and $x_{-s} \equiv \{x_i\}_{i\notin s}$, where $s \subseteq \mathcal{N}$ is an arbitrary coalition of players.

3 Review of Coalition-Stable Equilibria in Stage Games

In this section, I briefly summarize the results presented in Chung (2004). Section 3.1 presents a general critique on theories of rational behavior in games, and motivates the principles behind the concept of coalition-stable equilibria. Section 3.2 gives a formal definition of the solution concept, and highlights its key properties.

3.1 Motivation and A General Critique

Most equilibrium theories in games implicitly assume that players can coordinate their behavior by meaningful pre-play communication. In particular, the Nash theory can be understood as presenting a necessary condition for any reasonable definition of an *agreement* in the pre-play communication process.[13] However, it is also well known that some Nash equilibria are more plausible than others as outcomes of strategic play. Consider, for example, the following simple 2×2 game:

	B_1	B_2
A_1	3, 3	0, 0
A_2	0, 0	1, 1

Figure 1: A Common-Interest Coordination Game

Suppose the common prior of both players is focused on (A_2, B_2). If pre-play communication is modelled as many stages of cheap-talk games, where on each stage players simultaneously announce their own intended strategy in the actual play of game. Then the Nash theory concludes that (A_2, B_2) is a plausible agreement between the players, since any announcement indicating a unilateral deviation seems to conflict with individual rationality, and it should not be believed. Therefore, in the actual play of game, thinking that her opponent is going to play action 2, each

[12] Since Σ^0 is compact and $\alpha(\cdot)$ is continuous in the topology of pointwise convergence, $v_i(\cdot)$ is uniformly continuous.

[13] Otherwise, there is a players who is not playing her best response, and she should be able to convince her opponents that she wants to do something else.

player should best response and play action 2 herself, fulfilling the equilibrium prediction of (A_2, B_2).

Nevertheless, a simple forward-induction logic may break the above Nash argument. Suppose the row player announces she is going to play A_1. This is a "counterfactual" according to the Nash theory, since the theory predicts no rational player should have an incentive to make such an announcement.[14] But to conclude the row player is irrational conflicts with common knowledge of individual rationality. Hence, it is highly likely that the column player may search for a rational explanation of her opponent's behavior, before concluding the impossible.

In the present situation, it is only reasonable to interpret the counterfactual announcement as an invitation for the column player to announce B_1 in the next round of communication. First, (A_1, B_1) is a Nash equilibrium and hence is comfortable with a prediction of the existing rational theory. Second, if the column player believes in her opponent's announcement and acts according to the implicit invitation, there is no incentive for her opponent to cheat and do something else. Finally, it is reasonable for the row player to send and expect the column player to act on her invitation, since the outcome is mutually beneficial. Therefore, it seems unlikely that (A_2, B_2) would become an agreement in the pre-play communication stage and realize as an outcome of the game, even though it is a Nash equilibrium.

The above arguments indicate that the Nash theory, or any theory of rational behavior in general, may give rise to "counterfactuals" that are globally destabilizing. Sophisticated rational players should be able to exploit and manipulate counterfactuals of a given rational theory to their own strategic advantage. Hence not every prediction of a rational theory may be regarded as a plausible outcome of play, if the prediction is checked against the implications of counterfactuals produced by the theory.

The concept of *coalition-stable* equilibrium attempts to address the problems posed by counterfactuals by employing an iterated forward-induction argument. It represents a theory of rational behavior that is "complete", in the sense that common knowledge of the theory would not undermine its predictions through the counterfactuals it produces. In other words, no rational player can expect to gain by attempting counterfactuals produced by the theory developed here. Moreover, as evident from the above example, combining forward-induction logic with pre-play communication allows players to correlate their strategies, and exercise a form of coalitional reasoning. Hence for the rest of development in this paper, I can treat the present theory as a description of certain equilibrium behavior that is stable against some form of self-enforcing coalitional deviations, even though the motivation and the formal analysis of the theory is strictly non-cooperative.[15]

[14] Farrell (1993) called this type of counterfactual in the cheap-talk games a *neologism*.

[15] Interested readers are strongly encouraged to refer to Chung (2004) for a formal treatment of the theory.

3.2 Definition and Properties of Equilibrium

This section presents a construction of the set of coalition-stable equilibria in stage games, and highlights key properties of equilibrium. Loosely speaking, the set of equilibria can be constructed by the following successive refinement procedure. Let Σ^p denote the set of strategy profiles that is immune to self-enforcing deviations by coalitions of size less than or equal to p. By definition, $\Sigma^0 \supseteq \Sigma^1 \supseteq ... \supseteq \Sigma^n$, where Σ^n is the set of coalition-stable equilibria in an n-player game. For $p = 1, ..., n$, given that Σ^{p-1} has been previously defined, I can construct Σ^p as the intersection over sets of unimprovable strategy profiles in Σ^{p-1}, among coalitions of size p.[16] Notice that the first step of this iterative algorithm gives the set of Nash equilibria. Hence by construction, the theory of coalition-stable equilibrium refines the Nash theory using coalitional incentive in a natural way.[17]

To present a formal definition of the theory, first we need to develop a couple of mathematical notations. Given a subset of players $s \subseteq \mathcal{N}$ with size denoted by $|s|$, I can define a strict partial ordering $\succ_s$ over the space of feasible strategy profiles:[18] for all $\widetilde{\sigma}, \sigma \in \Sigma^0$,

$$\widetilde{\sigma} \succ_s \sigma \Leftrightarrow \begin{cases} u_i(\widetilde{\sigma}_s, \sigma_{-s}) > u_i(\sigma) \quad \forall i \in s \\ \widetilde{\sigma}_{-s} = \sigma_{-s} \end{cases}$$

Hence $\widetilde{\sigma} \succ_s \sigma$ if and only if $\widetilde{\sigma}_s$ can be interpreted as a profitable *coordinated* deviation by every player in s from σ, taking as fixed the strategic choices of players in the complement coalition $-s$. Given the partial ordering $\succ_s$ defined above, and a closed subset $\widetilde{\Sigma}$ of feasible strategy profiles, we say $\sigma \in \widetilde{\Sigma}$ is unimprovable in $\widetilde{\Sigma}$ by coalition s, if there does not exist any $\widetilde{\sigma} \in \widetilde{\Sigma}$ such that $\widetilde{\sigma} \succ_s \sigma$. Let $K_s(\widetilde{\Sigma})$ represents the set of unimprovable strategy profiles in $\widetilde{\Sigma}$ for coalition s. Mathematically, we have $K_s(\widetilde{\Sigma}) \equiv \{\sigma \in \widetilde{\Sigma} : \nexists \widetilde{\sigma} \in \widetilde{\Sigma},\ \widetilde{\sigma} \succ_s \sigma\}$.[19]

We are ready for a formal definition of equilibrium. It consists of a sequence of intermediate definitions $\{\Sigma^p\}_{p=1}^n$, representing successive refinements of Nash as an equilibrium theory of rational behavior. When $p = 1$, define $\Sigma^1 = \bigcap_{i \in \mathcal{N}} K_i(\Sigma^0)$, hence Σ^1 coincides with the set of Nash equilibria. Suppose Σ^m has been defined for $m = \{1, 2, ..., p-1\}$, set $\Sigma^p \equiv \bigcap_{s,|s|=p} K_s(\Sigma^{p-1})$.

[16] A strategy profile σ is unimprovable in Σ^p by a coalition s if, taking the choices of the complement coalition as given, there does not exists a joint-deviation by s that is strictly beneficial for every member of s, and the resulting strategy profile $\tilde{\sigma}$ is in Σ^p.

[17] Readers are again reminded that the phrase "coalitional incentive" only has descriptive contents. It is a *consequence* of applying forward-induction reasoning in the pre-play communication process. The *primitives* consist only of common knowledge of individual rationality and common prior over the space of strategy profiles. Hence the spirit of analysis is strictly non-cooperative. Please refer to section 3.1 and Chung (2004) for further discussion.

[18] A strict partial ordering is a binary relation $\succ$ that is asymmetric and transitive, that is, $a \succ b \Rightarrow b \nsucc a$, and $a \succ b$ & $b \succ c \Rightarrow a \succ c$, respectively.

[19] In other words, $K_s(\widetilde{\Sigma})$ is the mathematical core of the abstract system $(\widetilde{\Sigma}, \succ_s)$ in Greenberg (1989). Moreover, since $\succ_s$ is a strict partial ordering, it also coincides with the von Neuman/Morgenstern stable set of $(\widetilde{\Sigma}, \succ_s)$ defined in Von-Neumann and Morgenstern (1944).

Definition 1. Σ^n *is the set of coalition-stable equilibria in a n-player stage game.*

Beyond the literal interpretation that Σ^n is a refinement of Nash using incentives of coalitions of increasing sizes,[20] as hinted by the informal discussion in section 3.1, readers should be aware that each step of refinement from Σ^{p-1} to Σ^p also corresponds to one application of forward-induction logic, which refines the rational theory at the previous stage by deleting those predictions of the theory that are unstable, when checked against the implications of counterfactuals produced by the theory.[21] Hence the current solution concept may be best viewed as a positive theory of strategic play, in an environment where players can coordinate their expectations in an open pre-play communication process, and they attempt to explain every announcement by each player using the simplest theory consistent with common knowledge of individual rationality.[22]

Chung (2004) showed that the equilibrium correspondence is closed, and under weak technical conditions it is also non-empty. I restate an existence theorem presented there.[23]

Theorem 1 (Existence in Stage Games). *If Σ^0 is a nonempty, convex and compact topological vector space and the payoff profile $u(\cdot)$ is quasi-concave, then Σ^p is nonempty $\forall p \leq n$.*[24]

As an important corollary, since payoff profile $u(\cdot)$ is linear and hence quasi-concave when mixed-strategies are considered, Theorem 1 implies generic existence of coalition-stable equilibria. This is both a surprising and a desirable property of the equilibrium theory, since it is well-known that existing solution concepts that refine Nash using coalitional incentives, such as *strong Nash equilibria* proposed by Selten (1975) and *coalition-proof equilibria* proposed by Bernheim et al. (1987), have difficulties with existence in generic normal-form stage games.

4 Equilibria in Repeated Games

This section extends the theory of coalition-stable equilibria to situations where players are engaged in repeated interactions. Besides the initial pre-play communication, I assume players also have opportunities to communicate and revise their strategies at the beginning of each round of actual play. Section 4.1 uses several simple two-period examples to highlight the strategic complications

[20] We say a profitable joint-deviation from $\widetilde{\sigma}$ to σ by a coalition of size p is self-enforcing if and only if $\sigma \in \Sigma^{p-1}$.

[21] Successive refinement by incentives of coalitions of increasing sizes corresponds to the progressively sophisticated ways that rational players can manipulate counterfactuals in a given rational theory to their strategic advantage, since every coalitional proposal can be viewed as a chain of individual proposals in the pre-play communication process. Details are presented in Chung (2004).

[22] Similar behavioral assumptions also appear in the formulation of extensive-form rationality in Pearce (1984) and Battigalli (1997).

[23] The proof is reproduced in the Appendix.

[24] A payoff profile $u(\cdot) = \{u_i(\cdot)\}_{i\in\mathcal{N}}$ is *quasi-concave* if and only if $\forall\sigma, \tilde{\sigma} \in \Sigma^0$ and $\forall\lambda \in [0,1]$, $u(\lambda\sigma + (1-\lambda)\tilde{\sigma}) \geq u(\sigma) \wedge u(\tilde{\sigma}) = \{min[u_i(\sigma), u_i(\tilde{\sigma})]\}_{i\in\mathcal{N}}$.

introduced by the possibility of renegotiation. Section 4.2 introduces the concept of dynamic coalitions and presents a formal definition of equilibrium.

4.1 Simple Two-Period Examples

This section uses several two-period examples to illustrate how the possibility of renegotiation would change the strategic consideration of the players, and hence the solution concept used to predict the outcomes of the game. For simplicity, we consider only pure-strategy equilibria and assume no discounting in this section.

The first example concerns with the issue of dynamic consistency at the collective level. Consider the extended Prisoner's Dilemma game in figure 2.[25] There are two Nash equilibria, $(D1, D1)$ and $(D2, D2)$, in the stage game. Imagine a single repetition of this game. If the players can meet and coordinate their strategies *only* before the play of the game, then there is a perfect Nash equilibrium in which players initially cooperate (C, C), and play $(D1, D1)$ in the terminal period, with any first period deviation punished by reversion to $(D2, D2)$. This strategy profile gives each player a total payoff of 5. Moreover, since this equilibrium is Pareto-efficient among the class of Nash equilibria, it is also coalition-stable in the normal-form representation of this two-stage game.[26]

		Player 2		
		C	$D1$	$D2$
	C	3, 3	0, 4	0, 0
Player 1	$D1$	4, 0	2, 2	0, 0
	$D2$	0, 0	0, 0	1, 1

Figure 2: Extended Prisoner's Dilemma Game

If players can reconvene and reconsider their options after play in the first period, then the strategy profile described above is no longer immune to strategic manipulation. According to our discussion in section 3, the prescription to play $(D2, D2)$ in the second period is highly implausible, since players can use the communication opportunity before the second period to coordinate on a mutually preferable outcome $(D1, D1)$. Knowing the "punishment" to revert to $(D2, D2)$ is incredible, a player will deviate even in the first period, rendering cooperation in the first period untenable. Hence in the presence of renegotiation, the only reasonable prediction of this game is to play $(D1, D1)$ in both periods, which gives a total payoff of only 4 to each player. In the present example, renegotiation *decreases* ex-ante utility of each player, because players cannot *commit* to

[25] This example is borrowed by Table 1 of Bernheim and Ray (1989).

[26] For 2-player normal-form games, the concept of coalition-stable equilibria coincides with the solution concept of Pareto-undominated Nash.

a more severe off-equilibrium punishment ex-post.

	B_1	B_2
A_1	3, 1	0, 0
A_2	0, 0	1, 3

Figure 3: Battle of Sexes

The second example illustrates how re-opening communication before play at each stage help to discipline the forward-induction logic across different periods. Consider a single repetition of the Battle of Sexes game in figure 3. Suppose players originally agree to play $(A1, B1)$ in both periods. If players *cannot* communicate after play in the first period, then the column player (she) may have an incentive to deviate to $B2$ in the first period, even if she is rational. It is because if she knows that the row player (he) uses a forward-induction logic to interpret her deviation, then she can effectively use the deviation to signal her intention to play $B2$ in the second period, and also her expectation that the row player will cooperate by playing $A2$. Even if the row player prefers their original agreement to this new suggestion by the column player, he no longer has any opportunity to communicate to the column player and reiterates his intention to adhere to the old agreement. Hence the row player may submit to this signal sent by his opponent, and best responds to it by playing $A2$ in the second period, which in turn justifies the first-period deviation by the column player.

The above forward-induction argument is, however, fragile if players can communicate after play in the first period. This essentially turns the implicit one-way communication channel into an explicit two-way pre-play communication process in the subgame following the first-period moves. More importantly, it takes away the last-mover's advantage of the column player. Upon observing the first-period deviation by the column player, the row player can now reiterate his adherence to their original agreement and announce his intention to play $A1$ in the second period. Such an announcement is credible, since it is consistent with their original agreement, and since there is no reason why the row player should sacrifice his own interest for the interest of the column player. Hence in this case, the possibility of further communication before the second stage undermines the signalling effect of deviations in the first stage, and discourages the column player from first-period deviation. In other words, by arguing that players should adhere to their previous agreement in the subgame unless there is a mutually beneficial adjustment, we let the backward-induction logic overrides the forward-induction logic for intertemporal agreements.[27]

The third example challenges the conventional wisdom that renegotiation leads to agreements whose payoffs are Pareto-frontier of some admissible set.[28] Consider a single repetition of the game

[27] See van Damme (1989) for a similar advocation.

[28] This requirement is labelled as *weak renegotiation-proofness* by Farrell and Maskin (1989) and as *internal con-*

		Player 2		
		L	M	R
	L	5, 5	0, 0	0, 0
Player 1	M	0, 0	4, 4	2, 6
	R	0, 0	6, 2	0, 0

Figure 4: A Simple Game Challenging WRP

in figure 4. The Nash equilibria of the stage game are (L, L), (M, R) and (R, M), all of them are also coalition-stable in the stage game. Obviously, the strategy profile σ that specifies playing (L, L) in both periods is a perfect Nash equilibrium. Moreover, we can verify that $\{(M, M), (L, L)\}$ is also a perfect Nash equilibrium outcome supported by the following strategy profile $\tilde{\sigma}$:

$$\text{First Period:} \quad \text{play } (M, M)$$

$$\text{Second Period:} \quad \text{play} \begin{cases} (L, L) & \text{if no player deviates} \\ (M, R) & \text{if player 1 deviates alone} \\ (R, M) & \text{if player 2 deviates alone or jointly with player 1} \end{cases}$$

Since payoffs associate with σ Pareto dominates those of $\tilde{\sigma}$, the conventional wisdom suggests that players would collectively renounce $\tilde{\sigma}$ in favor of σ, if they can communicate and coordinate their strategies prior to the play of the game.

However, this argument by the conventional wisdom presumes that players can make binding commitment to switch from their original agreement $\tilde{\sigma}$ to the new agreement σ. Equivalently, it assumes players cannot reconvene and reconsider their strategies after play in period one. Otherwise, player 1 may not want to go along with σ in the second period, since adherence to σ only gives her a second-period payoff of 5, while the original agreement $\tilde{\sigma}$ promises her a payoff of 6 under such contingency. Furthermore, knowing that player 1 would not participate in the second-period deviation, player 2 would not deviate in the first period, since doing so yields her a total payoff of $5 + 2 = 7$, while the original agreement $\tilde{\sigma}$ gives her a total payoff of $4 + 5 = 9 > 7$. In the current scenario, renegotiation fails to increase ex-ante utility of both players, because players cannot *commit* to a particular course of action, and because they cannot prevent themselves from further reneging their current agreement in future.

4.2 Dynamic Coalitions and Definition of Equilibria

This section introduces the idea of dynamic coalitions, which embeds our intuition given in section 4.1, and extends our construction of coalition-stable equilibria to repeated games with discounting. I'll take as primitive the successive refinement procedure described in section 3, and modifies the construction in a minimal way to incorporate additional strategic considerations introduced by re-

sistency by Bernheim and Ray (1989).

peated interactions and negotiations.

The notion of dynamic coalitions incorporates the idea that players may be able to formulate an elaborate intertemporal scheme to deviate from an existing agreement. Moreover, since players' behavior is dynamically consistent, and since no binding commitment to any particular course of action is available, every side-agreement to deviate must give strict incentive to those who participate, at every contingency when the participant's action is called upon.

To describe a mutually beneficial deviation by a subset of players $s \in 2^{\mathcal{N}}$ at a particular contingency h_t,[29] I generalize the strict partial ordering defined in section 3.2 to the following: for all $\widetilde{\sigma}, \sigma \in \Sigma^0$,

$$\widetilde{\sigma} \succ_s^{h_t} \sigma \Leftrightarrow \begin{cases} U_i(\alpha((\widetilde{\sigma}_s, \sigma_{-s}), h_t)) > U_i(\alpha(\sigma, h_t)) \quad \forall i \in s \\ \widetilde{\sigma}_{-s}|_{h_t} = \sigma_{-s}|_{h_t} \end{cases}$$

Definition 2. *A dynamic coalition S is a sequence of history-dependent functions $\{S_t\}_{t=1}^T$ such that $S_t : H_t \to 2^{\mathcal{N}}$ $\forall t$.*

Let $\mathcal{C}$ be the space of non-empty dynamic coalitions.[30] Given a dynamic coalition $S \in \mathcal{C}$, I can construct a strict partial ordering $\succ_S$ corresponding to S by defining

$$\widetilde{\sigma} \succ_S \sigma \Leftrightarrow \widetilde{\sigma} \succ_{S_t(h_t)}^{h_t} \sigma \quad \forall h_t \in H_t, \forall t < \infty^{31}$$

Intuitively, $\widetilde{\sigma} \succ_S \sigma$ if and only if at every finite history h_t, $\widetilde{\sigma}|_{h_t}$ is a profitable *coordinated* deviation from $\sigma|_{h_t}$ for *all* players in coalition $s_t = S_t(h_t)$, taking as fixed strategies of players in the complement coalition $-s_t$. Hence $\widetilde{\sigma}$ fits in our description as an elaborate intertemporal scheme to deviate from an existing agreement σ, since at every contingency h_t, the players who are called upon to act by the scheme, $S_t(h_t)$, have a strict incentive to carry on the prescribed deviation.

We say $\widetilde{\sigma}$ is a *subgame-preferable* deviation from σ if $\exists S \in \mathcal{C}$ such that $\widetilde{\sigma} \succ_S \sigma$. Given the partial ordering $\succ_S$ constructed above, and a closed subset $\widetilde{\Sigma}$ of feasible strategy profiles, I can define $\sigma \in \widetilde{\Sigma}$ to be an unimprovable strategy profile in $\widetilde{\Sigma}$ for dynamic coalition S, if there does not exist any $\widetilde{\sigma} \in \widetilde{\Sigma}$ such that $\widetilde{\sigma} \succ_S \sigma$. Again, I'll denote by $K_S(\widetilde{\Sigma})$ the set of unimprovable strategy profiles in $\widetilde{\Sigma}$ for dynamic coalition S.[32]

The construction of coalition-stable equilibria in repeated games parallels its construction

[29]Notice that mathematically, the notion $s \in 2^{\mathcal{N}}$ includes the possibility that s is an empty set of coalition.

[30]$S \in \mathcal{C}$ if and only if $\exists h_t \in H_t$ such that $S_t(h_t) \neq \emptyset$.

[31]When $S_t(h_t) = \emptyset$, $\widetilde{\sigma} \succ_{S_t(h_t)}^{h_t} \sigma$ if and only if $\widetilde{\sigma}|_{h_t} = \sigma|_{h_t}$.

[32]Mathematically, $K_S(\widetilde{\Sigma}) \equiv \{\sigma \in \widetilde{\Sigma} : \nexists \widetilde{\sigma} \in \widetilde{\Sigma}, \ \widetilde{\sigma} \succ_S \sigma\}$ is the unique von Neuman/Morgenstern stable set of system $(\widetilde{\Sigma}, \succ_S)$, and inherits its mathematical properties such as compactness and non-emptiness.

for stage games. It also consists of a sequence of intermediate definitions $\{\Sigma^p\}_{p=1}^n$, representing successive refinements of perfect Nash as an equilibrium theory of rational behavior. Let $|S| \equiv |\bigcup_{h_t \in H_t, \forall t} S_t(h_t)|$ denote the size of a dynamic coalition S, and define $\mathcal{C}(p) \equiv \{S \in \mathcal{C} : |S| = p\}$ to be the set size-p dynamic coalitions. It is straight forward to verify that the set of subgame-perfect Nash strategy profiles can be represented by $\Sigma^1 = \bigcap_{i \in \mathcal{C}(1)} K_i(\Sigma^0)$.[33] Moreover, suppose Σ^m has been previously defined for $m = \{1, 2, ..., p-1\}$, I can then set $\Sigma^p \equiv \bigcap_{s \in \mathcal{C}(p)} K_s(\Sigma^{p-1})$.

Definition 3. Σ^n *is the set of coalition-stable equilibria in a n-player repeated game.*

First note that since the definition of equilibria is non-recursive, it applies equally well to both finitely and infinitely repeated games. Moreover, when attention is restricted to 2-player repeated games, this definition of equilibrium shares many common intuition laid out in the renegotiation-proof literature:[34] *a perfect Nash equilibrium σ is coalition-stable, if there does not exist another perfect Nash equilibrium $\tilde{\sigma}$, such that after some contingency h_t, players would jointly prefer to renounce σ for $\tilde{\sigma}$.* Finally, because the construction of equilibria in repeated games parallels its construction in stage games, it shares the same intuition that motivates this solution concept as previously discussed in section 3. Hence the solution concept in repeated games may be best viewed as a positive theory of strategic play, in an environment where players can openly communicate prior to each stage of action, and that players can effectively coordinate their behavior via counterfactuals implied by the simplest rational theory.

5 Recursive Characterization in Infinitely Repeated Games

From our definition of equilibrium and subgame-preferable deviations, it is obvious that for finitely repeated games, we can use the standard backward-induction procedure and characterize the set of equilibria recursively starting from the last period. However, this approach fails in infinitely repeated games, since there is no last period. Nevertheless, in this section, I show the set of equilibria has a surprisingly simple intertemporal structure, and its payoff set can be characterized recursively using an iterative procedure extending the algorithm pioneered by Abreu et al. (1990).

Since our construction of equilibria involves a sequence of intermediate definitions $\{\Sigma^p\}_{p=1}^n$, from sections 5.1 to 5.3, I'll take Σ^{p-1} as given and focus on developing a procedure to obtain Σ^p from Σ^{p-1}. Section 5.4 makes use of these developments and proposes an iterative algorithm to fully characterize the equilibrium payoff set. Section 5.5 further simplifies equilibrium charac-

[33]Sketch of proof: ($\Rightarrow$) Suppose $\sigma \notin \cap_{i \in \mathcal{C}(1)} K_i(\Sigma^0)$, then $\exists j \in \mathcal{C}(1)$ such that $\sigma \notin K_j(\Sigma^0)$. Hence $\exists \hat{\sigma} \in \Sigma^0$ such that $\hat{\sigma} \succ_j \sigma$, i.e. $\exists h_t$ such that $U_j(\alpha((\hat{\sigma}_j, \sigma_{-j}), h_t)) > U_j(\alpha(\sigma, h_t))$. Hence σ is not subgame perfect. ($\Leftarrow$) Suppose σ is not subgame perfect, then $\exists h_t$ such that $U_j(\alpha((\hat{\sigma}_j, \sigma_{-j}), h_t)) > U_j(\alpha(\sigma, h_t))$, and hence $\exists j \in \mathcal{C}(1), \hat{\sigma} \in \Sigma^0$ such that $\hat{\sigma} \succ_j \sigma$. Obviously, $\sigma \notin K_j(\Sigma^0) \Rightarrow \sigma \notin \cap_{i \in \mathcal{C}(1)} K_i(\Sigma^0)$.

[34]See for example Farrell and Maskin (1989), Bernheim and Ray (1989) and Abreu et al. (1993).

terization by optimal penal codes. Section 5.6 uses the recursive characterization to prove existence.

5.1 Equivalence Between One-Step Deviations and Finite-Step Deviations

This section marks the first step in simplifying our characterization of equilibrium, and prepares ourself for the recursive procedures developed in later sections. In particular, I show that to refine Σ^p from Σ^{p-1}, it suffice to check one-step and infinite-step deviations from size-p coalitions, because a version of "no-gain-from-one-shot-deviation" principle holds for all finite-step coalitional moves.

The proof of this version of "no-gain-from-one-shot-deviation" principle follows directly from our definition of subgame-preferable deviations and the usual intuition. Loosely speaking, for a group of players to participate in an elaborate intertemporal scheme of deviation from an existing agreement, we require each participant has a strict incentive to carry out the scheme, at every contingency when her action is called upon. If this intertemporal scheme of deviation involves only finitely-many step, then there is a last group of players who participate in the scheme, and each of them must strictly prefers to deviate at the last step than to revert to the original arrangement. Since participants at the "last step" essentially carry out an one-step deviation, immunity to one-step deviations is sufficient for immunity to every finite-step of deviation.

To formally develop the above intuition, we need some new mathematical notations. Given a strategy profile $\sigma \in \Sigma^0$, let $\mathcal{U}^{p,k}(\sigma)$ denote the set of subgame-preferable deviations from σ carried out by some size-p dynamic coalitions that are active for at-most k periods.[35] Define $\Sigma^{p,k} \equiv \{\sigma \in \Sigma^{p-1} : \Sigma^{p-1} \cap \mathcal{U}^{p,k}(\sigma) = \emptyset\}$. In other words, $\Sigma^{p,k}$ is the set of strategy profiles in Σ^{p-1} that are immune to subgame-preferable deviations in Σ^{p-1} by all size-p dynamic coalitions that are active for at-most k periods. By construction, we have $\Sigma^p = \Sigma^{p,\infty} \subseteq \Sigma^{p,k}\ \forall k < \infty$.

Proposition 1. $\Sigma^{p,1} = \Sigma^{p,k} \quad \forall\ 1 \leq k < \infty$

It is important to note that Proposition 1 is invalid when $k = \infty$. When considerations are restricted to the set of unilateral deviations ($p = 1$), the standard approach is to approximate every infinite deviation by its τ-period truncation for some large finite τ. This can be done because, with discounting, the payoff from every infinite-period deviation can be arbitrarily well approximated by its finite-period counterpart, and hence this finite-period truncation is still a profitable unilateral deviation for the individual player. Hence the standard approach shows Proposition 1 can be extended to $k = \infty$ for the set of unilateral deviations.

The standard argument, however, does not work when the deviation involves more than a single player. To see why, consider an infinite repetition of the Prisoner's Dilemma game in figure 5.

[35] Hence $\mathcal{U}^{p,k}(\sigma) = \{\widetilde{\sigma} \in \Sigma^0 : \exists S \in \mathcal{C}(p),\ \tau \geq 0\ s.t.\ \widetilde{\sigma} \succ_S \sigma,\ \widetilde{\sigma}_t = \sigma_t\ \forall t < \tau \text{ and } S_{k+t}(h_{k+t}) = \emptyset\ \forall h_{k+t},\ t \geq \tau\}$.

	C	D
C	3, 3	0, 4
D	4, 0	1, 1

Figure 5: A Standard Prisoner's Dilemma Game

When players are sufficiently patient ($\delta \approx 1$), the strategy profile $\tilde{\sigma}$ that specifies playing (C, D) in odd periods and (D, C) in even periods, is a subgame-preferable deviation from the strategy profile σ, that corresponds to the infinite repetition of static Nash (D, D).[36] Nevertheless, no finite-step truncation of $\tilde{\sigma}$ is a subgame-preferable deviation from σ, since the player who plays C at the last step necessarily wants to play D instead. As a result of this potential "discontinuity at infinity", we'll first move on to develop a generalization of the powerful decomposition techniques of Abreu et al. (1990) in section 5.2, and then come back for a more general proof of "no-gain-from-one-shot-deviation" principle in section 5.3.

5.2 Self-Generation and Factorization

Since every continuation of a strategy profile that is immune to finite-step deviations is itself immune to finite-step deviations, we can recursively characterize the set of payoffs associated with $\Sigma^{p,1}$ in the spirit of Abreu et al. (1990). Let the set of payoffs associated with Σ^{p-1} and $\Sigma^{p,1}$ be denoted by V^{p-1} and $V^{p,1}$, respectively.[37] I begin with the construction of a set-valued operator $B_p(\cdot)$ that captures the relationship between the set of promise utilities, and the set of payoffs associated with admissible strategies with these promises.

Definition 4. *Given an action profile $q \in A$, and a promise function $\rho : A \to \Re^n$, an action-promise pair (q, ρ) is **p-admissible** w.r.t. a closed set $W \subseteq V^{p-1}$ if and only if*

1. $\rho(q) \in W \ \forall q \in A$
2. $\forall s \subseteq \mathcal{N},\ |s| \leq p$, *and* $\forall \tilde{q}_s \in \{A_i\}_{i \in s}$ *s.t.* $(1-\delta)u(\tilde{q}_s, q_{-s}) + \delta\rho(\tilde{q}_s, q_{-s}) \in V^{p-1}$,

   $\exists i \in s$ *s.t.* $(1-\delta)u_i(q) + \delta\rho_i(q) \geq (1-\delta)u_i(\tilde{q}_s, q_{-s}) + \delta\rho_i(\tilde{q}_s, q_{-s})$

Let $E_i(q, \rho) \equiv (1-\delta)u_i(q) + \delta\rho_i(q)$ be the payoff for player i associated with the pair (q, ρ), and use $E(q, \rho) = \{E_i(q, \rho)\}_{i \in \mathcal{N}}$ to denote the vector of payoff profile for all players. I can define the set-valued operator $B_p(W) \equiv \{E(q, \rho) : (q, \rho) \text{ is } p\text{-}admissible \text{ w.r.t. } W\}$[38].

Definition 5. *$W \subset \Re^n$ is p-**self-generating** if and only if $W \subseteq B_p(W)$.*

[36] The payoffs corresponds to repeated static Nash σ is $(1, 1)\ \forall t$, while the payoffs from $\tilde{\sigma}$ is approximately $(2, 2)\ \forall t$.

[37] Hence $V^{p-1} = v\{v(\sigma) : \sigma \in \Sigma^{p-1}\}$ and $V^{p,1} = \{v(\sigma) : \sigma \in \Sigma^{p,1}\}$.

[38] When $p = 0$, we define $B_0(W) \equiv \{E(q, \rho) : \rho(\tilde{q}) \in W\ \forall q, \tilde{q} \in A\}$.

Our set-valued operator is a generalization of the $B(\cdot)$ operator defined in Abreu et al. (1990),[39] and hence it has similar mathematical properties such as monotonicity and upper hemicontinuity.[40] Finally, the following lemmas show that we can extend their self-generation and factorization techniques, and relate each p-self-generating set to payoff profiles of a strategy in $\Sigma^{p,1}$, and associate the largest fixed point of $B_p(\cdot)$ with the set $V^{p,1}$.

Proposition 2 (p-Self-Generation). *If W is p-self-generating, then $B_p(W) \subseteq V^{p,1}$.*

Proposition 3 (p-Factorization). $V^{p,1} = B_p(V^{p,1})$

Intuition and proof of these propositions follow closely to the development in Abreu et al. (1990) and are left to the Appendix.

5.3 A Generalized No-Gain-From-One-Shot-Deviation Principle

We are now ready to show that the "no-gain-from-one-shot-deviation" principle also applies to infinite-step coalitional deviation, and hence we can formally associate the largest fixed point of $B_p(\cdot)$ with the set of payoff profiles obtained from Σ^p. The intuition of the proof is as follows: if $\sigma \in \Sigma^{p,1}$ is immune to every self-enforcing finite-step deviations, but $\hat{\sigma} \in \Sigma^{p-1}$ is an infinite-step subgame-preferable deviation from σ, then $\hat{\sigma}$ is necessarily a Ponzi-scheme in promise utilities along the deviation path. Hence no such $\hat{\sigma}$ is feasible.

Theorem 2. $\Sigma^p = \Sigma^{p,1}$

Sketch of Proof Given $\sigma \in \Sigma^{p,1}$ and suppose $\exists \hat{\sigma} \in \Sigma^{p-1}$ representing an infinite-step subgame-preferable deviation by some size-p dynamic coalition $S \in \mathcal{C}(p)$. W.l.o.g., we can assume the deviation starts at $t = 0$, i.e. $\hat{\sigma}_0 \neq \sigma_0$.[41] Then for all histories $\hat{h}_t = (\hat{q}_0, \hat{q}_1, ..., \hat{q}_{t-1})$ along the deviation path induced by $\hat{\sigma}$,[42] by definition of subgame-preferable deviation we have $v_{s_t}(\lambda^t(\hat{\sigma}|_{\hat{h}_t})) > v_{s_t}(\lambda^t(\sigma|_{\hat{h}_t}))\ \forall t$, where $s_t = S_t(\hat{h}_t)$, and $\lambda^t(\hat{\sigma}|_{\hat{h}_t})$ and $\lambda^t(\sigma|_{\hat{h}_t})$ are the continuation strategies at $\hat{h}_t$ for $\hat{\sigma}$ and σ, respectively.[43] Since $\sigma \in \Sigma^{p,1}$, no τ-step deviation along $\hat{h}_t$ is profitable. Hence

$$\forall t,\ \exists i \in s_t \quad v_i(\lambda^t(\sigma|_{\hat{h}_t})) \geq (1-\delta)[\sum_{\nu=0}^{\tau-1} \delta^\nu u_i(\hat{q}^{t+\nu})] + \delta^\tau v_i(\lambda^{t+\nu}(\sigma|_{\hat{h}_{t+\nu}}))$$

[39] $B_p(\cdot)$ is equivalent to the $B(\cdot)$ operator in Abreu et al. (1990) when $p = 1$.

[40] See Lemmas 2 and 3 in the Appendix.

[41] If the infinite-step deviation starts at $h_t \neq h_0$, then we can consider the continuation strategy profile of $\hat{\sigma}$ at h_t, $\lambda^t(\hat{\sigma}|_{h_t}) \in \Sigma^{p-1}$, as an infinite-step deviation from the continuation strategy profile of σ at h_t, $\lambda^t(\sigma|_{h_t}) \in \Sigma^{p,1}$.

[42] $\{\hat{h}_t\}_{t=0}^{\infty}$ can be obtained recursively by $\hat{h}_1 = \hat{q}_0 = \hat{\sigma}_0$ and $\hat{h}_{t+1} = (\hat{h}_t, \hat{q}_t)$ where $\hat{q}_t = \hat{\sigma}_t\ (h_t)\ \forall t \geq 0$.

[43] $\lambda^t(\sigma) \equiv \{\sigma_\tau\}_{\tau=t}^{\infty}$ is the shift operator that deletes the first t elements of σ.

Since $v(\lambda^t(\hat{\sigma}|_{\hat{h}_t})) - v(\lambda^t(\sigma|_{\hat{h}_t})) = (1-\delta)[\sum_{\nu=0}^{\tau-1}\delta^\nu(u(\hat{q}^{t+\nu}) - v(\lambda^t(\sigma|_{\hat{h}_t})))] + \delta^\tau[v(\lambda^{t+\tau}(\sigma|_{\hat{h}_{t+\tau}})) - v(\lambda^t(\sigma|_{\hat{h}_t}))] + \delta^\tau[v(\lambda^{t+\tau}(\hat{\sigma}|_{\hat{h}_{t+\tau}})) - v(\lambda^{t+\tau}(\sigma|_{\hat{h}_{t+\tau}}))]$, we have:

$$\begin{aligned}
& \max_{i\in s_t}\left\{\delta^\tau[v_i(\lambda^{t+\tau}(\hat{\sigma}|_{\hat{h}_{t+\tau}})) - v_i(\lambda^{t+\tau}(\sigma|_{\hat{h}_{t+\tau}}))] - [v_i(\lambda^t(\hat{\sigma}|_{\hat{h}_t})) - v_i(\lambda^t(\sigma|_{\hat{h}_t}))]\right\} \\
= & \max_{i\in s_t}\left\{(1-\delta)[\textstyle\sum_{\nu=0}^{\tau-1}\delta^\nu(v_i(\lambda^t(\sigma|_{\hat{h}_t})) - u_i(\hat{q}^{t+\nu}))] + \delta^\tau[v_i(\lambda^t(\sigma|_{\hat{h}_t})) - v_i(\lambda^{t+\tau}(\sigma|_{\hat{h}_{t+\tau}}))]\right\} \\
= & \max_{i\in s_t}\left\{v_i(\lambda^t(\sigma|_{\hat{h}_t})) - (1-\delta)[\textstyle\sum_{\nu=0}^{\tau-1}\delta^\nu u_i(\hat{q}^{t+\nu})] - \delta^\tau v_i(\lambda^{t+\nu}(\sigma|_{\hat{h}_{t+\nu}}))\right\} \geq 0
\end{aligned}$$

Hence $\forall t, \tau < \infty,\ \exists i \in s_t$ such that

$$v_i(\lambda^{t+\tau}(\hat{\sigma}|_{\hat{h}_{t+\tau}})) - v_i(\lambda^{t+\tau}(\sigma|_{\hat{h}_{t+\tau}})) \geq \frac{1}{\delta^\tau}\left[v_i(\lambda^t(\hat{\sigma}|_{\hat{h}_t})) - v_i(\lambda^t(\sigma|_{\hat{h}_t}))\right]$$

In particular, for $t = 0$, since $\min_{i\in s_0}\{v_i(\lambda^0(\hat{\sigma}|_{\hat{h}_0})) - v_i(\lambda^0(\sigma|_{\hat{h}_0}))\} = \min_{i\in s_0}\{v_i(\hat{\sigma}) - v_i(\sigma)\} > 0$ and $\delta < 1$, $\exists \hat{\tau} < \infty$ such that $\max_{i\in s_0} v_i(\lambda^{\hat{\tau}}(\hat{\sigma}|_{\hat{h}_{\hat{\tau}}})) > M\ \forall M < \infty$. But this is impossible since the set of feasible utility V^0 is compact and hence bounded in each dimension.

5.4 A Recursive Characterization of Equilibrium Payoffs

This section summarizes the developments in sections 5.1 - 5.3 by providing an iterative algorithm which completely characterizes the equilibrium payoff set V^n. Therefore, we can use the tracing procedure outlined in the proof of Proposition 2, and recover every equilibrium outcome of a given infinitely repeated game.

Our algorithm consists of n stages of set-valued iterations, where on each stage we take the limit set of the previous stage, and iterate "downwards" until convergence using a fixed $B_p(\cdot)$ operator. Hence each stage of our iterations is akin to the computational procedure described in Abreu et al. (1990). To start the chain of iterations on each step, we need the following proposition:

Proposition 4. $V^{p-1} \supseteq B_p(V^{p-1}) \supseteq V^{p,1} \quad \forall p \geq 1$

Theorem 3. *Define* $W_0^p = V^{p-1}$ *and* $W_k^p = B_p(W_{k-1}^p)$. *Then* $W_k^p \supseteq W_{k+1}^p$ *and* $V^p = \cap_k W_k^p = \lim_{k\to\infty} W_k^p$. *Moreover,* $\lim_{k\to\infty} W_k^p$ *is compact.*

Theorem 3 suggests the following algorithm to characterize the equilibrium payoff set V^n: Start with V^0 and iterate downwards using $B_1(\cdot)$ until convergence to approximate V^1. Given that V^{p-1} has been previously characterized, compute V^p by iterating downwards from V^{p-1} using $B_p(\cdot)$ until convergence, and successively calculate V^p for each p until $p = n$.

5.5 Optimal Penal Codes

The set-valued iteration procedure outlined in section 5.4 may be difficult to implement in practice. Fortunately, in this section I show that a version of Abreu (1988)'s optimal penal codes exists for our solution concept. Hence every equilibrium outcome can be easily characterized, given these penal codes have been found. In section 6, I'll show how to support collusion in an infinitely repeated Cournot duopoly game with renegotiation using the idea of optimal penal codes.

As in Abreu (1988), each optimal penal code α^i corresponds to one of the worst equilibrium outcome paths for each player $i \in \mathcal{N}$. Notice that we do *not* need to tailor an optimal penal code for any coalition of players. The insight is that to discourage a coalition of players from deviating, it suffices to punish one of its member very harshly immediately after a group deviation has been observed. Again, the stick-and-carrot structure of these codes guarantees the punished player has an incentive to voluntarily participate in her own punishment.

Let $\rho^{i,p} \in argmin\{w_i | w \in V^{p,1}\}$ denote the payoff-profile associated with the most severe punishment in $\Sigma^{p,1}$ for player $i \in \mathcal{N}$. We first present a general proposition that leads to the existence of optimal penal codes.

Proposition 5. *$w \in V^p$ if and only if there exists (q, ρ) p-admissible w.r.t. V^p such that $(1-\delta)u(q) + \delta\rho(q) = w$ and $\rho(\tilde{q}) \in \{\rho^{i,p}\}_{i \in \mathcal{N}}$ $\forall \tilde{q} \neq q$.*

Proof We only need to prove the only-if part. Since $\Sigma^p = \Sigma^{p,1} \Rightarrow V^p = B_p(V^p)$, given $w \in V^p, \exists (q, \rho)$ p-admissible w.r.t. V^p such that $(1-\delta)u(q) + \delta\rho(q) = w$. Hence $\forall \tilde{q}_s, |s| \leq p$, with $(1-\delta)u(\tilde{q}_s, q_{-s}) + \delta\rho(\tilde{q}_s, q_{-s}) \in V^{p-1}$, there exists $i \in s$ such that $(1-\delta)u_i(\tilde{q}_s, q_{-s}) + \delta\rho_i(\tilde{q}_s, q_{-s}) \leq (1-\delta)u_i(q) + \delta\rho_i(q) = E_i(q, \rho)$. Since $\rho^{i,p} \leq \rho_i(\tilde{q}_s, q_{-s})$ $\forall i \in s$, we have $\max_{i \in s}\{E_i(q, \rho) - (1-\delta)u_i(\tilde{q}_s, q_{-s}) - \delta\rho_i^{i,p}\} \geq 0$. Define $\hat{\rho}$ by $\hat{\rho}(q) = \rho(q)$, and $\forall \tilde{q} = (\tilde{q}_s, q_{-s}) \neq q$,[44] $\hat{\rho}(\tilde{q}_s, q_{-s}) = \rho^{\hat{i},p}$ where $\hat{i} \in argmax_{i \in s}\{E_i(q, \rho) - (1-\delta)u_i(\tilde{q}_s, q_{-s}) - \delta\rho_i^{i,p}\}$. It is straight forward to verify that $(q, \hat{\rho})$ is p-admissible w.r.t. V^p and $E(q, \hat{\rho}) = w$.

Let $\sigma^i \in \Sigma^n$ with $v(\sigma^i) = \rho^{i,n}$ be the worst equilibrium strategy profile for player i, and let $\alpha^i \equiv \alpha(\sigma^i)$ be the outcome path associated with σ^i. Proposition 5 implies that we can use $\{\alpha^i\}_{i \in \mathcal{N}}$ as the set of optimal penal codes to support every equilibrium outcome. Moreover, these codes are "simple" in the sense that the *set* of penal codes is history-independent, and that every deviation is punished by restarting one of the codes in this set.

Notice that, however, our penal codes are different from Abreu (1988)'s original proposal in one crucial aspect: the choice of punishment may depend on the deviation. In particular, the player

[44]When $\tilde{q} \neq q$ has multiple representations as $\tilde{q} = (\tilde{q}_s, q_{-s})$, pick the s with the smallest size $|s|$.

$i \in s$ who is punished after deviation $\tilde{q}_s$ from an equilibrium outcome path $\alpha = (q_0, q_1, ...)$ is chosen by maximizing over $i \in s$:

$$U_i(\alpha) - (1-\delta)u_i(\tilde{q}_s, q_{0,-s}) - \delta U_i(\alpha^i)$$

Hence the player selected to be punished is the player who has the most to lose from the coalitional deviation, and therefore is "weakest link" of the coalition.[45,46]

5.6 Existence of Equilibrium

This section uses the recursive characterization developed in section 5.4 to prove existence of equilibrium in infinitely repeated games. The approach to existence proof is complicated by the fact that, in general, infinite repetition of an equilibrium in a stage game need *not* be an equilibrium in the corresponding infinitely repeated game. Fortunately, I can show that existence of equilibrium in stage game is sufficient for existence in infinitely repeated game.[47] Moreover, since Theorem 1 implies generic existence of equilibrium in stage games, existence of our solution concept is guaranteed in behavior strategies in repeated games.[48]

In this section, let $\{\Sigma^p_{stage}\}^n_{p=0}$ denote the corresponding concepts $\{\Sigma^p\}^n_{p=0}$ in the stage game. Our existence proof is implied by the following proposition:

Theorem 4. $\Sigma^p_{stage} \neq \emptyset \Rightarrow V^p \neq \emptyset$

Idea of Proof Notice that $q \in \Sigma^p_{stage} \Rightarrow (1-\delta)u(q) + \delta w \in B_p(\{w\})\ \forall w \in V^{p-1}$, since players are induced to play the stage game if promise utilities do not vary with current action. Since V^p is the limit of the decreasing sequence $\{B^m_p(V^{p-1})\}^\infty_{m=0}$ and $B_p(\cdot)$ is monotonic, we have $B^{m+1}_p(V^{p-1}) = B_p\left(B^m_p(V^{p-1})\right) \supseteq \bigcup_{w \in B^m_p(V^{p-1})} B_p(\{w\}) \neq \emptyset\ \forall m \Rightarrow V^p \neq \emptyset.$

6 An Example of Repeated Cournot Duopoly with Renegotiation

This section uses an example of repeated Cournot duopoly game to illustrate the idea of optimal penal codes developed in section 5.5. In particular, I'll demonstrate how these penal codes, which essentially represent a *stick-and-carrot* strategy on the individual level but a *divide-and-conquer* strategy on the group level, can be used to support collusion even in the presence of renegotiation. It is also interesting to note that since infinite repetition of Cournot-Nash is the only perfect Nash

[45] I thank Jonathan Levin for suggesting this descriptive phrase to me.

[46] DeMarzo (1992) used a similar idea to support his solution concept for finitely repeated games.

[47] With minor modifications, our approach can also be used to prove existence in finitely repeated games.

[48] Please also refer to footnote 10 for a remark on the observability of players' private randomization device in repeated games.

equilibrium that is also "weakly renegotiation-proof",[49] our theory gives a very different prediction compared to the solution concepts developed by Farrell and Maskin (1989) and Bernheim and Ray (1989), who both predict that collusion is impossible in the current environment if renegotiation is allowed.

Consider an infinite repetition of the two-player stage game in Figure 6.[50] Roughly speaking, the game can be thought of as a symmetric, discrete, quantity-setting duopoly game played by two firms, each of them may choose a low (L), medium (M), or high (H) output level. The unique Nash equilibrium of the stage game is (M, M). Let δ be the discount factor. It could be easily verified Cournot-Nash reversion can be used to support the "collusive" outcome $\{(L, L), (L, L), ...\}$ if and only if $\delta \geq 4/7$.

		Firm 2		
		L	M	H
	L	10, 10	5, 14	0, 6
Firm 1	M	14, 5	7, 7	-5, -2
	H	6, 0	-2, -5	-15, -15

Figure 6: Repeated Cournot Duopoly with Renegotiation

Let $\delta = 1/2$, so that Cournot-Nash reversion is not severe enough to support the "collusive" outcome. Nevertheless, Abreu (1988) showed that $\{(L, L), (L, L), ...\}$ can still be supported by the following pair of optimal penal codes:

$$\underline{Q}^1 = \{(M, H), (L, M), (L, M), ...\}, \quad \underline{Q}^2 = \{(H, M), (M, L), (M, L), ...\}^{51}$$

His idea is that the collusive outcome can be attained if each firm $\{i = 1, 2\}$ is threatened with the worst perfect Nash equilibrium outcome $\underline{Q}^i$ in case of an unilateral deviation. Moreover, the stick-and-carrot structure of these codes guarantees each firm has an individual incentive to follow equilibrium recommendation at all histories. However, since these codes hurt both the innocent and the guilty, and joint-deviations are not punished, firms jointly prefer to skip the "stick" phase and jump directly to the "carrot" phase, whenever these penal codes are called for. It can easily be checked that the resulting paths $\underline{\hat{Q}}^1 = \{(L, M), (L, M), ...\}$ and $\underline{\hat{Q}}^2 = \{(M, L), (M, L), ...\}$ cannot serve as penal codes, since the guilty does not have an incentive to participate in his own punishment.

[49] A perfect Nash equilibrium is "weakly renegotiation-proof", if it is not possible to find two different histories, h_t and h'_t, such that the continuation utilities given by the equilibrium strategy profile at these two contingencies are Pareto-ranked. Bernheim and Ray (1989) use the term "internal consistency" to represent essentially the same idea.

[50] This game is a modified version of the example given in Abreu (1988).

[51] Notice that $\underline{Q}^1$ and $\underline{Q}^2$ are the worst subgame-perfect Nash equilibrium outcomes for firm 1 and firm 2, respectively, with the corresponding equilibrium payoff profile of $(0, 6)$ and $(6, 0)$.

In the presence of renegotiation, I show that collusion can still be supported, if we adopt a strategy of "divide-and-conquer" on the group level. This can be done by adopting a pair of less aggressive simple penal codes, and by making joint-deviations punishable:

$$\underline{Q}^{1*} = \{(L,H),(L,H),(L,M),(L,M),...\}, \quad \underline{Q}^{2*} = \{(H,L),(H,L),(M,L),(M,L),...\}$$

As in Abreu (1988), I require every unilateral deviation by firm i be punished by restarting $\underline{Q}^{i*}$. In addition, every joint-deviation along the path of $\underline{Q}^{i*}$ is punished by restarting $\underline{Q}^{j*}$ with $j \neq i$.[52] Notice the way I punish joint-deviations is first to divide the interest of the group on such deviations, and then to conquer them by punishing at the "weakest link" of the group, i.e. the firm who has the most to lose if it participates in the group deviation. Finally, these penal codes are optimal in the sense that $\underline{Q}^{i*}$ corresponds to the worst coalition-stable equilibrium outcome for firm $i = 1, 2$.

We can easily verify that no firm has an incentive to deviate from collusion if $\{\underline{Q}^{i*}\}_{i=1,2}$ are used as punishments. To see if firms have an incentive to deviate from either $\underline{Q}^{1*}$ or $\underline{Q}^{2*}$, Theorem 2 shows that we only need to check for one-shot deviations. First note that all unilateral deviations can be checked in the usual manner. Moreover, the "stick-and-carrot" structure of each of these penal codes implies joint-deviation from $\underline{Q}^{i*}$ reduces promise utility of firm $j \neq i$. Hence we only need to check the joint-deviation to (M, M), the only group deviation that is jointly profitable in the current period.

Suppose equilibrium recommends restarting $\underline{Q}^{1*}$. If both firms follow the recommendation, the payoff-pair they received is $(\frac{5}{4}, 8)$. If both firms jointly deviate in the first period, then the payoff-pair they receive is $\frac{1}{2} \cdot (7,7) + \frac{1}{2} \cdot (8, \frac{5}{4}) = (\frac{15}{2}, \frac{33}{8})$. Since firm 2 is made worse off by this joint-deviation ($\frac{33}{8} < 8$), she will not participate in the first period. Now suppose both firms follow equilibrium recommendation in the first period, and they enter the second period with a promised utility-pair of $(\frac{5}{2}, 10)$. If these firms jointly deviate in the second period, the utility-pair they receive is again $(\frac{15}{2}, \frac{33}{8})$. But then obviously firm 2 would not like to participate since ($\frac{33}{8} < 10$). Moreover, there is no mutually beneficial one-shot joint-deviation from $\{(L,M),(L,M),...\}$, and hence $\underline{Q}^{1*}$ is immune to coalitional deviation starting from the third period. Finally, since the game is symmetric, arguments for "renegotiation-proofness" of $\underline{Q}^{2*}$ is analogous.

7 Conclusion

This section concludes by relating our theory to other solution concepts in the literature. We start by comparing the theory to two benchmark notions of renegotiation-proofness – the concept of *weak*

[52] In this example, I can punish all joint-deviations from $\{(L,L),(L,L),...\}$ by either $\underline{Q}^{1*}$ or $\underline{Q}^{2*}$. In general, however, the punishment path chosen can depends on the particular joint-deviation, even though the set of punishment paths remains the same over all contingencies. Please refer to section 5.3 for a more precise statement.

renegotiation-proofness proposed by Farrell and Maskin (1989) and Bernheim and Ray (1989),[53] and a different idea of renegotiation-proofness advocated by Pearce (1987) and Abreu et al. (1993). As we shall see below, while each of these benchmark notions represents quite an extreme but different view over the power of current-period players in influencing the course of strategic play, the view adopted by our theory is quite balanced, and is guided completely by the dynamically consistent behavior of players. As such, our theory contributes to the literature by reconciling the difference between these two benchmark notions of renegotiation-proofness, and by presenting a coherent framework to analyze the problem of renegotiation.

The notion of *weak renegotiation-proofness*, first introduced by Farrell and Maskin (1989) and Bernheim and Ray (1989), represents a rather optimistic view concerning the renegotiation power of current-period players in repeated games. This notion requires that no two points in the equilibrium payoff set should be Pareto-ranked.[54] The intuition is that if both σ and $\tilde{\sigma}$ are candidates for an equilibrium, but the payoffs associated with σ strictly Pareto-dominate those associated with $\tilde{\sigma}$, then collective rationality of players should lead them to jointly abolish $\tilde{\sigma}$ in favor of σ, even if $\tilde{\sigma}$ is recommended. Hence $\tilde{\sigma}$ cannot be a candidate for a "renegotiation-proof" equilibrium.

While the argument for *weak renegotiation-proofness* is valid in stage games, as we have discussed in the third example (Figure 4) of section 4.1, it is in general invalid even in two-period games. The problem is that the joint-deviation from $\tilde{\sigma}$ to σ may require participation of players in future periods. Since dynamic consistency implies players cannot commit themselves to a particular course of action, and since the usual Pareto-criterion compares only the utilities of players in the current period but ignores the promises to players in the future periods, Pareto-dominance in current-period utilities is in general not a sufficient condition for renegotiation. Hence players may not be able to renegotiate to a better outcome in the current period, because the current agreement to deviate may be further reneged in future.

As a consequence, even though *weak renegotiation-proofness* is an intuitively appealing requirement, it implies players can make intertemporal binding commitment when renegotiating a side-contract to deviate. Since an equilibrium agreement is required to be self-enforcing precisely because we assume no commitment technology is available, the requirement of *weak renegotiation-proofness* may not be appropriate in general.

On the other end of the spectrum, the notion of renegotiation-proofness adopted by the theories of Pearce (1987) and Abreu et al. (1993) represents a rather pessimistic view over the renegotiation

[53] Bernheim and Ray (1989) labels the same idea *internal consistency*. This idea is also closely related to the concept of *internal stability* advocated by Asheim (1991), and the idea of *internally renegotiation-proofness* proposed by Ray (1994). Please refer to footnote 49 for a definition of *weak renegotiation-proofness*.

[54] When the game is only finitely repeated, this criterion is sometimes called *Pareto dominance refinement*, which says no two points in the equilibrium payoff set in any subgame are Pareto-ranked.

power of players.[55] Embedded in their theories is the idea that, for a joint-deviation from one perfect Nash equilibrium $\tilde{\sigma}$ to another perfect Nash equilibrium σ to be credible, not only should such a deviation give a strictly higher payoff to every players in the current period, but it is also required that under no circumstance would this deviation result in a lower payoff to any player in future. In essence, their theories give all future-period players a veto power to reject every joint-deviation proposed by players at present, even though the joint-deviation may be one-shot and therefore require no participation of players in future periods. Obviously, their requirement for credible renegotiation may be too stringent.

Our theory adopts a balanced view on the renegotiation power of players, and is guided completely by concerns over their dynamically consistent behavior. Our theory neither allows players to commit to strategic choices made in future, nor let future-period players who are uninvolved in a deviation to block the strategic changes made by players at present. As such, our theory escapes the difficulties associated existing benchmark notions of renegotiation-proofness discussed above, and in a sense embraces different conflicting views in the literature in a single coherent framework.

Our theory is related to the concept of *Perfectly Coalition-Proof Nash equilibrium* of Bernheim et al. (1987), which described strategic outcomes in finitely repeated games where coalitional plans to deviate could be kept secret from other players. The theory is also related to the concept of *Sustainable Social Norms* proposed by DeMarzo (1992), who extended *Strong Nash equilibrium* of Aumann (1959) to repeated games by considering strategy profiles that are not "sequentially blocked", a concept which is closely related to our notion of subgame-preferable deviations. Finally, the idea developed here is also related to Osborne (1990), who considered how the signalling effect of an individual deviation can be used to refine predictions in finitely repeated games.

The theory developed here proposed a solution to the problem of renegotiation, in an environment with no nature moves and with no private information. Extending the current theory to a more general environment is a challenging but rewarding exercise and will be taken in future works. Recently, Ambrus (2002) and Ambrus (2003) extends the concept of rationalizability of Bernheim (1984) and Pearce (1984) to incorporate incentives of coalitions. It would also be interesting to explore the relationship between our equilibrium concept and his rationalizability concept in future works.

[55] Readers should note that the exact definition of "renegotiation-proof" equilibrium in Abreu et al. (1993) is slightly different from that of Pearce (1987), even though the general idea is the same.

References

Abreu, Dilip, "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, March 1988, *56* (2), 383 – 396.

——, **David Pearce, and Ennio Stacchetti**, "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, September 1990, *58* (5), 1041 – 1063.

——, ——, **and** ——, "Renegotiation and Symmetry in Repeated Games," *Journal of Economic Theory*, 1993, *60*, 217 – 240.

Ambrus, Attila, "Coalitional Rationality," 2002.

——, "Dynamic Coalitional Agreements – coalitional rationalizability in muti-stage games," 2003.

Asheim, Geir, "Extending Renegotiation-Proofness to Infinite Horizon Games," *Games and Economic Behavior*, 1991, *3*, 278 – 294.

Aumann, Robert, "Acceptable Points in General Cooperative n-Person Games," in "Contributions to the Theory of Games IV," Princeton University Press, 1959.

Battigalli, Pierpaolo, "On Rationalizability in Extensive Games," *Journal of Economic Theory*, 1997, *74*, 40 – 61.

Bergin, James and W. Bentley MacLeod, "Efficiency and Renegotiation in Repeated Games," *Journal of Economic Theory*, 1993, *61*, 42 –73.

Bernheim, B. Douglas and Debraj Ray, "Collective Dynamic Consistency in Repeated Games," *Games and Economic Behavior*, 1989, *1*, 295 – 326.

——, **Bezalel Peleg, and Michael D. Whinston**, "Coaltion-Proof Nash Equilibria I. Concepts," *Journal of Economic Theory*, June 1987, *42* (1), 1 – 12.

Bernheim, Douglas, "Rationalizable Strategic Behavior," *Econometrica*, 1984, *52*, 1007 – 1028.

Chung, Anthony Fai-Tong, "Preplay Communication, Iterated Forward Induction, and Coalition-Stable Equilibria," Mimeo. Department of Economics, Stanford University, 2004.

DeMarzo, Peter, "Coalitions, Leadership, and Social Norms: The Power of Suggestion in Games," *Games and Economic Behavior*, 1992, *4*, 72 – 100.

Farrell, Joseph, "Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior*, 1993, *5*, 514 – 431.

—— **and Eric Maskin**, "Renegotiation in Repeated Games," *Games and Economic Behavior*, 1989, *1*, 327 – 360.

Greenberg, Joseph, "Deriving Strong and Coalition-Proof Nash Equilibria from an Abstract System," *Journal of Economic Theory*, 1989, *49*, 195 – 202.

Harris, Christopher, "Existence and Characterization of Perfect Equilibrium in Games of Perfect Information," *Econometrica*, 1985, *53* (3), 613 – 628.

Kaplan, G., "Sophisticated Outcomes and Coalitional Stability," Master's thesis, Deparment of Statistics, Tel Aviv University 1992.

Milgrom, Paul and John Roberts, "Coalition-Proofness and Correlation with Arbitrary Communication Possibilities," *Games and Economic Behavior*, 1996, *17*, 113 – 128.

Osborne, Martin J., "Signaling, Forward Induction, and Stability in Finitely Repeated Games," *Journal of Economic Theory*, 1990, *50*, 22 – 36.

Pearce, David, "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 1984, *52*, 1029 – 1050.

——, "Renegotiation-Proof Equilibria: Collective Rationality and Intertemporal Cooperation," Discussion Paper 855, Cowles Foundation December 1987.

Ray, Debraj, "Internally Renegotiation-Proof Equilibrium Sets: Limit Behavior with Low Discounting," *Games and Economic Behavior*, 1994, *6*, 162 – 177.

Selten, Robert, "Reexamination of Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 1975, *4*, 25 – 55.

van Damme, Eric, "Stable Equilibria and Forward Induction," *Journal of Economic Theory*, 1989, *48*, 476 – 496.

Von-Neumann, John and Oskar Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944.

A Existence of Equilibrium in Stage Games

Given $\sigma, \tilde{\sigma} \in \Sigma^0$, we say $\sigma \succ^p \tilde{\sigma}$ if and only if $\exists s \subseteq \mathcal{N},\ |s| = p$, such that $\sigma \succ_s \tilde{\sigma}$. Define $\triangle^p(\sigma) \equiv \{\tilde{\sigma} \in \Sigma^0 : \sigma \succ^p \tilde{\sigma}\}$ and $\triangle_s(\sigma) \equiv \{\tilde{\sigma} \in \Sigma^0 : \sigma \succ_s \tilde{\sigma}\}$.[56] Obviously, $\triangle^p(\sigma) = \bigcup_{s,|s|=p} \triangle_s(\sigma)$ and define the nonempty compact correspondence $\psi^p : \Sigma^{p-1} \to co(\Sigma^{p-1})$ by $\psi^p(\sigma) = co(\Sigma^{p-1}) \setminus \triangle^p(\sigma)$.[57]

Lemma 1. *If Σ^0 is a nonempty, convex topological vector space and the payoff profile $u(\cdot)$ is quasi-concave, then i.) $\sigma \notin co\{\hat{\sigma} \in \Sigma^{p-1} : \hat{\sigma} \succ^p \sigma\}$, and ii.) ψ^p is a KKM correspondence, i.e. $co\{\sigma^1, ...\sigma^m\} \subseteq \bigcup_{i=1}^m \psi^p(\sigma^i)$ for every finite subset $\{\sigma^1, ...\sigma^m\}$ of Σ^{p-1}.*

Proof *i.)* Obviously $\sigma \not\succ^p \sigma$ since $\succ^p$ is irreflexive. Suppose $\exists \sigma^1, \sigma^2 \in \{\hat{\sigma} \in \Sigma^{p-1} : \hat{\sigma} \succ^p \sigma\}$ and $\lambda \in [0,1]$ such that $\sigma = \lambda\sigma^1 + (1-\lambda)\sigma^2$. This implies $\exists s \subseteq \mathcal{N}, |s| = p$ such that $\sigma^1 \succ_s \sigma$ and $\sigma^2 \succ_s \sigma$.[58] However, by quasi-concavity of $u(\cdot)$, $u_s(\lambda\sigma^1 + (1-\lambda)\sigma^2) \geq u_s(\sigma^1) \wedge u_s(\sigma^2) > u_s(\sigma) \Rightarrow \sigma \neq \lambda\sigma^1 + (1-\lambda)\sigma^2$, a contradiction. *ii.)* Fix an arbitrary finite subset $\{\sigma^1, ..., \sigma^m\}$ of Σ^{p-1} and pick $\tilde{\sigma} \in co\{\sigma^1, ..., \sigma^m\}$. Assume to the contrary that $\tilde{\sigma} \notin \bigcup_{i=1}^m \psi^p(\sigma^i)$. Then $\tilde{\sigma} \in co(\Sigma^{p-1}) / \bigcup_{i=1}^m \psi^p(\sigma^i) = \bigcap_{i=1}^m \triangle^p(\sigma^i) \Rightarrow \sigma^i \succ^p \tilde{\sigma}\ \forall i$. But $\tilde{\sigma} \in co\{\sigma^1, ..., \sigma^m\}$ implies $\tilde{\sigma} \in co\{\hat{\sigma} \in \Sigma^{p-1} : \hat{\sigma} \succ^p \sigma\}$, a contradiction.

Theorem 1 *If Σ^0 is a nonempty, convex and compact topological vector space and the payoff profile $u(\cdot)$ is quasi-concave, then Σ^p is nonempty $\forall p \leq n$.*

Proof It is sufficient to show Σ^{p-1} nonempty $\Rightarrow$ Σ^p nonempty given quasi-concavity of the payoff profile $u(\cdot)$. Fix an arbitrary $\hat{\sigma} \in \Sigma^{p-1}$ and define $\tilde{\psi}^p : co(\Sigma^{p-1}) \to co(\Sigma^{p-1})$ by setting $\tilde{\psi}^p(\hat{\sigma}) = \Sigma^{p-1} \cap \psi^p(\hat{\sigma})$, $\tilde{\psi}^p(\sigma) = \psi^p(\sigma)\ \forall \sigma \neq \hat{\sigma}, \sigma \in \Sigma^{p-1}$ and let $\tilde{\psi}^p(\sigma) = co(\Sigma^{p-1})\ \forall \sigma \in co(\Sigma^{p-1}) \setminus \Sigma^{p-1}$. Since ψ^p is a KKM correspondence by lemma 1, $\tilde{\psi}^p$ is a KKM correspondence by construction.[59] Moreover, $\succ_s$ is transitive $\forall s \subseteq \mathcal{N}$ implies $\Sigma^p = \bigcap_{\{s,|s|=p\}} \{\Sigma^{p-1} \backslash \triangle_s(\Sigma^{p-1})\} = \Sigma^{p-1} \backslash \triangle^p(\Sigma^{p-1}) = \bigcap_{\{\sigma \in \Sigma^{p-1}\}} \Sigma^{p-1} \backslash \triangle^p(\sigma) = \Sigma^{p-1} \cap \{\bigcap_{\{\sigma \in \Sigma^{p-1}\}} \psi^p(\sigma)\} = \bigcap_{\{\sigma \in co(\Sigma^{p-1})\}} \tilde{\psi}^p(\sigma)$. Hence we can prove non-emptiness of Σ^p by showing that $\{\tilde{\psi}^p(\sigma) : \sigma \in co(\Sigma^{p-1})\}$ is a collection of compact sets with finite-intersection property. Pick an arbitrary finite subset $\{\sigma^1, ..., \sigma^m\}$ of $co(\Sigma^{p-1})$. Let M be the finite dimensional space spanned by $\{\sigma^1, ..., \sigma^m\}$ and define $\{F_i\}_{1 \leq i \leq m}$ by $F_i = \tilde{\psi}^p(\sigma^i) \cap co\{\sigma^1, ..., \sigma^m\}$. Since $\tilde{\psi}^p$ is a KKM correspondence, $\{F_i\}_{1 \leq i \leq m}$ is a collection of compact subsets of M satisfying $co\{\sigma^i : i \in A\} \subseteq \cup_{i \in A} F_i$ for every subset of indexes $A \subseteq \{1, ..., m\}$. Therefore $\cap_{i=1}^m \tilde{\psi}^p(\sigma^i) \supseteq \cap_{i=1}^m F_i \neq \emptyset$ by KKM lemma. Since $\{\sigma^1, ..., \sigma^m\}$ is arbitrary, the collection $\{\tilde{\psi}^p(\sigma) : \sigma \in co(\Sigma^{p-1})\}$ has finite-intersection property which implies $\Sigma^p = \bigcap_{\{\sigma \in co(\Sigma^{p-1})\}} \tilde{\psi}^p(\sigma) \neq \emptyset$.

B Other Proofs

Let $\lambda^t(\sigma) \equiv \{\sigma_\tau\}_{\tau=t}^\infty$ be the shift operator that deletes the first t elements of σ, and write $\lambda^t(\widetilde{\Sigma}) = \bigcup_{\sigma \in \widetilde{\Sigma}} \lambda^t(\sigma)$. Since $G^\infty(\delta)$ is isomorphic to each of its own subgame, by construction we have $\lambda^t(\Sigma^p|_{h_t}) = \Sigma^p\ \forall h_t,\ 0 \leq p \leq n$, where $\Sigma^p|_{h_t} \equiv \{\sigma|_{h_t} : \sigma \in \Sigma^p\}$ is the restriction of Σ^p to history h_t.

[56] For every $\tilde{\Sigma} \subseteq \Sigma^0$, we denote $\triangle^p(\tilde{\Sigma}) = \cup_{\sigma \in \tilde{\Sigma}} \triangle^p(\sigma)$ and we define $\triangle_s(\tilde{\Sigma})$ analogously.

[57] Notice that $\sigma \in \psi^p(\sigma)$ since $\succ^p$ is irreflexive.

[58] If $\exists \sigma^1, \sigma^2 \in \Sigma^{p-1}$ with $\sigma^1 \succ_s \sigma$ and $\sigma^2 \succ_{\tilde{s}} \sigma$ but $s \neq \tilde{s}$, it can be easily shown that $\lambda\sigma^1_{\hat{s}} + (1-\lambda)\sigma^2_{\hat{s}} \neq \sigma_{\hat{s}}$, where $\hat{s} = (s \backslash \tilde{s}) \cup (\tilde{s} \backslash s)$.

[59] Notice that $\forall \sigma \in \Sigma^{p-1}$ with $\sigma \neq \hat{\sigma}$, $\tilde{\psi}^p(\hat{\sigma}) \cup \tilde{\psi}^p(\sigma) = (\Sigma^{p-1} \cap [\triangle^p(\hat{\sigma})]^c) \cup \psi^p(\sigma) = \{\Sigma^{p-1} \cup \psi^p(\sigma)\} \cap \{[\triangle^p(\hat{\sigma})]^c \cup \psi^p(\sigma)\} = \{co(\Sigma^{p-1}) \cup \psi^p(\sigma)\} \cap \{[\triangle^p(\hat{\sigma})]^c \cup \psi^p(\sigma)\} = \psi^p(\hat{\sigma}) \cup \psi^p(\sigma)$.

Proposition 1 $\Sigma^{p,1} = \Sigma^{p,k} \;\; \forall\, 1 \le k < \infty$

Proof Since $\mathcal{U}^{p,k}(\sigma) \supseteq \mathcal{U}^{p,1}(\sigma)\ \forall \sigma \in \Sigma^0, p \ge 1,\ \sigma \in \Sigma^{p,k} \Rightarrow \sigma \in \Sigma^{p,1}$. Hence $\Sigma^{p,k} \subseteq \Sigma^{p,1}$. Suppose $\sigma \notin \Sigma^{p,k}$, then $\exists \hat{\sigma} \in \Sigma^{p-1} \cap \mathcal{U}^{p,k}(\sigma)$. W.l.o.g. we can assume the chain of subgame-preferable deviations starts at $t = 0$. Therefore we have $\hat{\hat{\sigma}} \equiv \lambda^k(\hat{\sigma}|_{h_k}) \in \lambda^k(\Sigma^{p-1}|_{h_k}) \cap \lambda^k((\mathcal{U}^{p,k}(\sigma))|_{h_k}) = \Sigma^{p-1} \cap \mathcal{U}^{p,1}(\lambda^k(\sigma|_{h_k}))$. Finally, suppose $\sigma \in \Sigma^{p,1}$, then by definition of $\Sigma^{p,1}$ we have $\lambda^k(\sigma|_{h_k}) \in \Sigma^{p,1}$, but then $\hat{\hat{\sigma}} \in \Sigma^{p-1} \cap \mathcal{U}^{p,1}(\lambda^k(\sigma|_{h_k})) \neq \emptyset \Rightarrow \lambda^k(\sigma|_{h_k}) \notin \Sigma^{p,1}$.

Lemma 2. *Given $W_0, W_1 \subseteq V^{p-1}$. If $W_0 \supset W_1$, then $B_p(W_0) \supseteq B_p(W_1)$.*

Proof Obvious from the definition of $B_p(\cdot)$.

Lemma 3. *W is compact $\Rightarrow$ $B_p(W)$ is compact.*

Proof Consider the pair $(q, \rho) \in A \times A \times W,\ W \subseteq V^{p-1}$. Define $\mathcal{R}_w(V^{p-1}) \equiv \{(q,\rho) : E(q,\rho) \in V^{p-1}\}$. Given $s \subseteq \mathcal{N},\ |s| \le p$, construct a strict partial ordering $\gg_s$ on $\mathcal{R}_w(V^{p-1})$ as follows:

$$(q^1, \rho^1) \gg_s (q^2, \rho^2) \Leftrightarrow \begin{cases} E_i(q^1,\rho^1) > E_i(q^2,\rho^2) & \forall i \in s \\ q^1_{-s} = q^2_{-s},\ \rho^1 = \rho^2 \end{cases}$$

Let $\hat{K}_s \equiv \{(q,\rho) \in \mathcal{R}_w(V^{p-1}) : \nexists (\hat{q},\hat{\rho}) \in \mathcal{R}_w(V^{p-1}),\ (\hat{q},\hat{\rho}) \gg_s (q,\rho)\}$. Since both $A \times A \times W$ and V^{p-1} are compact, and $E(q,\rho) = (1-\delta)u(q) + \delta\rho(q)$ is continuous in (q,ρ), we have both $\mathcal{R}_w(V^{p-1})$ and $\hat{K}_s$ compact subsets of $A \times A \times W$. Moreover, (q,ρ) is p-admissible w.r.t. W if and only if $(q,\rho) \in \cap_{s,|s| \le p} \hat{K}_s$. Hence $\{(q,\rho) \in A \times A \times W : (q,\rho)$ is p-admissible w.r.t. $W\} = \cap_{s,|s|\le p} \hat{K}_s$ is compact. Therefore $B_p(W)$ is compact by continuity of $E(q,\rho)$.

Proposition 2 *If W is p-self-generating, then $B_p(W) \subseteq V^{p,1}$.*

Proof By the axiom of choice, $\forall w \in B_p(W)$, $\exists \hat{Q} : B_p(W) \to A$, and $\hat{P} : B_p(W) \to A \times W$ such that $(\hat{Q}(w), \hat{P}(w))$ is *p-admissible* w.r.t. W and $E(\hat{Q}(w), \hat{P}(w)) = w$. Hence for any $w \in W \subseteq B_p(W)$, we can construct $\sigma_w \in \Sigma^{p,1}$ with $v(\sigma_w) = w$ as follows. First, recursively define $\hat{P}^t_w : H_t \to \Re^n$ so that $\hat{P}^0_w = \hat{P}(w)$ and $\hat{P}^t_w(h_t) = \hat{P}(\hat{P}^{t-1}_w(h_{t-1}))(q_{t-1})\ \forall t \ge 1$. Since $\hat{P}(w) \in A \times W$ whenever $w \in B_p(W)$, and $W \subseteq B_p(W)$, by induction, we have $\hat{P}(w) \in A \times W \Rightarrow \hat{P}^t_w(h_t) \in W \subseteq B_p(W)\ \forall h_t \in H_t,\ t < \infty$. Let $\sigma_w = \{\sigma_{w,t}\}_{t=0}^{\infty}$ be defined by $\sigma_{w,0} = \hat{Q}(w),\ \sigma_{w,t} = \hat{Q}(\hat{P}^{t-1}_w(h_{t-1}))\ \forall t \ge 1$. By construction, $v(\sigma_w) = w$. Moreover, given $\{\hat{P}^t_w\}_{t=0}^{\infty}$, p-admissibility w.r.t. the p-self-generating set $W \subseteq V^{p-1}$ implies there does not exist profitable and self-enforcing one-step deviation $\hat{\sigma}_w \in \Sigma^{p-1}$ by any dynamic coalition of size $\le p$. Hence $\sigma_w \in \Sigma^{p,1}$. Since this is true for $\forall w \in B_p(W)$, therefore $B_p(W) \subseteq V^{p,1}$.

Proposition 3 $V^{p,1} = B_p(V^{p,1})$

Proof We only need to prove $V^{p,1} \subseteq B_p(V^{p,1})$. Given $w \in V^{p,1}$, $\exists \sigma \in \Sigma^{p,1}$ such that $v(\sigma) = w$. Let (q,ρ) be such that $q = \sigma_0,\ \rho(q) = v(\lambda^1(\sigma|_{h_1 = q}))$. Obviously, $E(q,\rho) = (1-\delta)u(q) + \delta\rho(q) = (1-\delta)u(\sigma_0) + \delta v(\lambda^1(\sigma|_{h_1=\sigma_0})) = v(\sigma) = w$. Also, $\sigma \in \Sigma^{p,1} \Rightarrow \lambda^1(\sigma|_{h_1}) \in \Sigma^{p,1}\ \forall h_1 \in H_1 \Rightarrow \rho(q) \in V^{p,1}\ \forall q$. Moreover, by definition of $\Sigma^{p,1}, (q,\rho)$ is p-admissible w.r.t. $V^{p,1}$. Hence $w \in B_p(V^{p,1})$.

Proposition 4 $V^{p-1} \supseteq B_p(V^{p-1}) \supseteq V^{p,1} \;\; \forall\, p \geq 1$

Proof When $p = 1$, it is obvious that $V^0 \supseteq B_1(V^0)$ since $V^0 = \{v(\sigma)|\sigma \in \Sigma\}$ is the set of payoff profiles from all feasible strategy profiles. When $p > 1$, given $w \in B_p(V^{p-1})$, $\exists (q^w, \rho^w)$ p-admissible w.r.t. V^{p-1} such that $(1-\delta)u(q^w) + \delta\rho^w(q^w) = w$. If $w \notin V^{p-1}$, since $\rho^w(q) \in V^{p-1}$ $\forall q \in A$, $\exists s$, $|s| \leq p-1$, and $\tilde{q}_s \in \{A_i\}_{i \in s}$ such that $(1-\delta)u(\tilde{q}_s, q^w_{-s}) + \delta\rho^w(\tilde{q}_s, q^w_{-s}) \in V^{p-1}$ and $E_i((\tilde{q}_s, q^w_{-s}), \rho^w) > E_i(q^w, \rho^w)$ $\forall i \in s$. This implies (q^w, ρ^w) is not p-admissible w.r.t. V^{p-1}, a contradiction. Hence $B_p(V^{p-1}) \subseteq V^{p-1}$ $\forall p \geq 1$. Moreover, since $V^{p-1} \supseteq V^{p,1}$, monotonicity of the $B_p(\cdot)$ operator implies $V^{p-1} \supseteq B_p(V^{p-1}) \supseteq B_p(V^{p,1}) = V^{p,1}$.

Theorem 3 *Define* $W^p_0 = V^{p-1}$ *and* $W^p_k = B_p(W^p_{k-1})$. *Then* $W^p_k \supseteq W^p_{k+1}$ *and* $V^p = \cap_k W^p_k = \lim_{k\to\infty} W^p_k$. *Moreover,* $\lim_{k\to\infty} W^p_k$ *is compact.*

Proof By iterating on the monotone operator $B_p(\cdot)$ and applying Proposition 4, we have

$$V^{p-1} \supseteq B_p(V^{p-1}) \supseteq ... \supseteq B^n_p(V^{p-1}) \supseteq ... \supseteq V^{p,1}$$

Then by definition of W^p_k, we have $W^p_k \supseteq W^p_{k+1}$ and $W^p_\infty \equiv \cap_k W^p_k \supseteq V^{p,1}$. Moreover, since $B_p(\cdot)$ maps a compact set to a compact set, W^p_k is compact $\forall k$, and therefore W^p_∞ is also compact. Also, because the graph of $B_p(\cdot)$ is compact, by closed graph theorem, $B_p(\cdot)$ is upper hemi-continuous, hence $W^p_\infty \subseteq B_p(W^p_\infty)$. Then by Proposition 2, $W^p_\infty \subseteq V^{p,1} \Rightarrow W^p_\infty = V^{p,1}$. Finally, Theorem 2 implies $V^{p,1} = V^p$, hence $W^p_\infty = V^p$.

Lemma 4. *Given* $q \in A$, $q \in \Sigma^p_{stage} \Leftrightarrow (1-\delta)u(q) + \delta w \in B_p(\{w\})$ $\forall w \in V^{p-1}$

Proof The lemma is obviously true when $p = 1$. Suppose the lemma is true for $p-1$, and fix some $w \in V^{p-1}$. If $q \in \Sigma^p_{stage}$, then $\forall s$, $|s| \leq p$, and $\forall \tilde{q}_s \in \{A_i\}_{i \in s}$ with $(\tilde{q}_s, q_{-s}) \in \Sigma^{p-1}_{stage}$, there exists $i \in s$ such that $u_i(q) \geq u_i(\tilde{q}_s, q_{-s})$. This implies $\forall s$, $|s| \leq p$, and $\forall \tilde{q}_s \in \{A_i\}_{i \in s}$ with $(1-\delta)u(\tilde{q}_s, q_{-s}) + \delta w \in B_{p-1}(\{w\})$, there exists $i \in s$ such that $(1-\delta)u_i(q) + \delta w_i \geq (1-\delta)u_i(\tilde{q}_s, q_{-s}) + \delta w_i$. Hence $(1-\delta)u(q) + \delta w \in B_p(\{w\})$. The proof in the other direction is completely analogous.

Theorem 4 $\Sigma^p_{stage} \neq \emptyset \Rightarrow V^p \neq \emptyset$

Proof Since repetition of a Nash equilibrium for the stage game is a subgame perfect Nash equilibrium for the infinitely repeated game, the statement is true when $p = 1$. Suppose the statement is true for $p-1$. Therefore $\Sigma^p_{stage} \neq \emptyset \Rightarrow \Sigma^{p-1}_{stage} \neq \emptyset \Rightarrow V^{p-1} \neq \emptyset$. Hence $B_p(\cdot)$ is well defined and by Theorem 3, $V^p = B^\infty_p(V^{p-1})$. Suppose $V^p = \emptyset$. Since the sequence $\{B^k_p(V^{p-1})\}^\infty_{k=0}$ consists of non-increasing compact sets having finite intersection property, $V^p = \emptyset \Rightarrow \exists m < \infty$ such that $B^m_p(V^{p-1}) \neq \emptyset$ but $B^{m+1}_p(V^{p-1}) = \emptyset$. However, by lemma 4, $\Sigma^p_{stage} \neq \emptyset \Rightarrow B_p(\{w\}) \neq \emptyset$ $\forall w \in V^{p-1}$. Hence monotonicity of $B_p(\cdot)$ implies $B^{m+1}_p(V^{p-1}) = B_p\left(B^m_p(V^{p-1})\right) \supseteq \bigcup_{w \in B^m_p(V^{p-1})} B_p(\{w\}) \neq \emptyset$ since $B^m_p(V^{p-1}) \subseteq V^{p-1}$, a contradiction. Hence $V^p \neq \emptyset$.