Commitment Through Renegotiation-Proof Contracts under Asymmetric Information*

Emanuele Gerratana
SIPA, Columbia University

Levent Koçkesen†
Koç University

March 27, 2013

Abstract

This paper characterizes equilibrium outcomes in extensive form games with incomplete information in which players sign renegotiable contracts with third-parties. Our aim is to understand whether renegotiation-proof third-party contracts can be used as commitment devices. We first characterize renegotiation-proof contracts and strategies for general extensive form games with incomplete information and then apply our results to two-stage games. If contracts are observable, then the second mover obtains her best possible payoff given that she plays a renegotiation-proof strategy and the first mover best responds. If contracts are unobservable, then this outcome is still an equilibrium but there are others. In fact a *"folk theorem"* type result holds: Any outcome in which the second mover best responds to the first mover's action on the equilibrium path and the first mover receives at least his "individually rational payoff", can be supported. We also apply our results to games with externalities and show that renegotiation-proofness imposes a very particular restriction in these games.

JEL Classification: C72, D80, L13.

Keywords: Third-Party Contracts, Commitment, Strategic Delegation, Renegotiation, Asymmetric Information, Renegotiation-Proofness, Entry-Deterrence.

*We would like to thank Alp Atakan, Emre Ozdenoren, Guofu Tan, seminar participants at GAMES 2012 World Congress (2012), Econometric Society European Meeting (2012), and Marmara University for helpful discussions. This research has been supported by TÜBİTAK Grant No. 106K317.

†Corresponding author: Department of Economics, Koç University, Rumelifeneri Yolu, Sariyer 34450, Istanbul, Turkey. Tel: +90 212 338-1354. E-mail: lkockesen@ku.edu.tr.

1 Introduction

Could an incumbent firm deter entry by contracting with third parties, such as a bank or a labor union? Could a central bank (or a government in a union) credibly commit to monetary policy (or fiscal policy) through a contract with the government (or supranational body)? More generally can contracts with third parties change the outcome of a game to the advantage of the contracting player? When contracts are non-renegotiable, the answer to this question is in general yes.[1] In fact, there are several "folk theorem" type results for different classes of games with *observable and non-renegotiable* third-party contracts.[2] The effects of *unobservable and non-renegotiable* third-party contracts are also well-understood: Nash equilibrium outcomes of a game with and without third-party contracts are identical (Katz (1991)). In fact, all (and only) Nash equilibrium outcomes of the original game can be supported as a sequential equilibrium outcome of the game with unobservable and non-renegotiable contracts (Koçkesen and Ok (2004) and Koçkesen (2007)). In this paper we seek an answer to this question for *renegotiable* contracts.[3]

More precisely, we analyze if an how renegotiation-proof third-party contracts change the equilibrium outcomes of extensive form games with incomplete information. In the main body of the paper we consider only two-player two-stage games where the second mover (player 2) has some payoff relevant private information. In what we call the *original game*, Nature moves first and determines the state of the world θ. After that, player 1 chooses an action a_1 without observing θ. Player 2 observes both θ and a_1, chooses a_2, and the game ends. Payoff function of player $i = 1,2$ is $u_i(a_1, a_2, \theta)$. Player 1's strategy is simply a choice of action a_1 whereas player 2's strategy is a function $b_2(a_1, \theta)$.

In the *game with contracts* we let player 2 sign a contract with a neutral third-party before the original game starts. A contract specifies transfers between player 2 and the third-party as a function of the contractible outcomes, which we assume to be the action choices of the two players, (a_1, a_2). The underlying and crucial assumption is that the private information of player 2 is not observable by any other player, including the third-party, and thus non-contractible. Given a contract f, the third-party's payoff is $f(a_1, a_2)$ whereas player 2's is $u_2(a_1, a_2, \theta) - f(a_1, a_2)$.

Our main objective is to understand the outcomes of the original game that can be supported in some equilibrium of the game with contracts. The first question that we need to answer is the type of strategies $b_2(a_1, \theta)$ that can be supported by a contract, i.e., incentive compatible strategies. Since contracts cannot depend on θ, incentive compatibility imposes some restrictions on b_2. In order to get a handle on these restrictions, we assume that player 2's payoff function exhibits increasing differences in (θ, a_2). It then follows that strategy b_2 is incentive compatible if and only if it is increasing in θ (This is Lemma 1 on page 7).

The second important step is to characterize the restrictions imposed by renegotiation. We model renegotiation as a game form: After player 1 moves, player 2 can make a renegotiation offer to the third-party, who knows a_1, but not θ, and can either accept the offer or reject it. We define renego-

[1] See, among many others, Vickers (1985), Fershtman and Judd (1987), Sklivas (1987), Koçkesen et.al. (2000), Brander and Lewis (1986), Bolton and Scharfstein (1990), Snyder (1996), Spencer and Brander (1983), Brander and Spencer (1985), Eaton and Grossman (1986), Walsh (1995).

[2] See Fershtman, Judd, and Kalai (1991), Polo and Tedeschi (2000), and Katz (2006).

[3] Prat and Rustichini (2003) and Jackson and Wilkie (2005) analyze related models in which players can write action contingent contracts before the game is played. Unlike the current paper, in these papers contractual relationships are not exclusive and the focus is on the efficiency properties of the equilibrium set. Also related is Bhaskar (2009), in which players need to pay a price to a supplier in order to play certain actions that are controlled by this supplier.

tiation-proof equilibrium as a perfect Bayesian equilibrium in which the equilibrium contract is not renegotiated after any (θ, a_1) and characterize the renegotiation-proof contracts and strategies (This is Theorem 1 on page 9). We also provide necessary and sufficient conditions for a strategy to be renegotiation-proof (Propositions 2 and 3).[4] These results generalize quite readily to arbitrary extensive form games with incomplete information where players are free to use mixed strategies, to environments in which the third-party is not neutral, and to stronger notions of renegotiation-proofness (See Section 4).

In Section 5 we present the implications of the above results in terms of the outcomes of the original game. We allow contracts to be observable or unobservable (by player 1) and renegotiable or non-renegotiable. We show that if contracts are observable, then player 2 can commit credibly to his Stackelberg payoff, i.e., the best payoff that he can achieve given that player 1 plays a best response. If the contracts are non-renegotiable, then the only restriction on the Stackelberg payoff is that player 2 uses an increasing strategy (Proposition 5). If contracts are renegotiable, then they also have to be renegotiation-proof (Proposition 7). We also show that these are the only outcomes that can be supported (Propositions 6 and 8). In other words, as long as one respects the restrictions imposed by incentive compatibility and renegotiation-proofness, contracts indeed serve as credible commitment devices.

We next consider unobservable contracts. We show that if contracts are non-renegotiable, then any Bayesian Nash equilibrium of the original game in which player 2's strategy is increasing can be supported (Proposition 9). In fact, we prove a folk theorem type result: any outcome $(a_1^*, a_2^*(\theta))$ of the original game in which $a_2^*(\theta)$ is a best response to a_1^* for each θ and player 1's payoff is at least as large as his "individually rational" payoff, can be supported (Corollary 1). Definition of individually rational payoff is different from the standard one in that player 2, in minimizing player 1's payoff, is restricted to using increasing strategies. Similar results hold for renegotiation-proof contracts except that in the definition of the individually rational payoff, player 2's strategy is restricted to be increasing and renegotiation-proof (Proposition 10 and Corollary 2). The upshot is that unobservable and renegotiation-proof contracts may still serve as commitment devices but unlike with observable contracts Stackelberg payoff is not the only equilibrium outcome anymore.

In Section 6 we provide two applications of our results. In one application we consider games with externalities, i.e., games in which player 1's payoff is increasing or decreasing in Player 2's action (Section 6.1). Suppose, for example, that player 1's payoff is increasing in a_2. If contracts are non-renegotiable, then player 2 can obtain a favorable outcome by punishing player 1 by playing the smallest a_2 whenever he plays an unfavorable action. Since a constant strategy is increasing, incentive compatibility does not bring any further restrictions on the outcomes that can be supported with non-renegotiable contracts. Renegotiation-proofness, on the other hand, imposes a very specific type of constraint on the kind of punishment player 2 can inflict upon player 1: the highest type of player 2 must play a best response while the other types could keep playing the smallest action (Corollary 3).[5] In other words, the additional restriction renegotiation-proofness brings about depends on the

[4] Our assumption that the third-party cannot observe θ during renegotiation is crucial. Otherwise, the result is trivial: One can only support the perfect Bayesian equilibria of the original game. This is because, if both a_1 and θ are common knowledge, then player 2 and the third-party would renegotiate away any strategy of player 2 that does not maximize the joint surplus, i.e., player 2's payoff in the original game.

[5] This is true when player 1's payoff is increasing in player 2's action. If his payoff is decreasing, then the harshest punishment player 2's can impose is to play the highest action for all types other than the lowest, who must play a best response.

probability of the highest type: If this probability is high, then renegotiation has a real bite, otherwise it does not.

The class of games with externalities is large and contains many economic models. The canonical example, of course, is the Stackelberg competition. We show that in this game, the follower firm indeed benefits from renegotiation-proof third-party contracts. This game can also be construed as an entry deterrence game, in which case we show that entry can always be deterred with non-renegotiable contracts but only under certain conditions with renegotiation-proof contracts, i.e., renegotiation has a real bite in these games.

RELATIONSHIP TO THE LITERATURE

The closest paper to ours is Dewatripont (1988), which analyzes an entry-deterrence game in which the incumbent signs a contract with a labor union before the game begins. A potential entrant observes the contract and then decides whether to enter or not. Renegotiation takes place after the entry decision is made, during which the union offers a new contract to the incumbent. The crucial assumption is that the incumbent has some payoff relevant private information during the renegotiation process. Dewatripont (1988) shows that commitment effects exist in such a model and may deter entry when contracts are publicly observable. We show that this commitment effect exists for unobservable contracts as well. Also, we analyze arbitrary two-stage games and hence can gauge the effects of third-party contracts in other interesting settings, for example in oligopoly models with price competition or credibly commitment to monetary and fiscal policy. Lastly, in our renegotiation protocol, the informed party makes the new contract offer, whereas in Dewatripont's, it is the uninformed party who makes the offer. This turns out to make a difference as we discuss in Section 6.1 (Proposition 11).

In a related paper, Gerratana and Koçkesen (2012) also study the effects of renegotiation-proof third-party contracts in two-stage games. However, that paper assumes that the original game is with perfect information whereas the current one assumes it is a game with incomplete information. Some aspects of the analyses of these two models are similar and use similar tools, namely theorems of the alternative. Indeed, results on renegotiation-proof contracts and strategies in Section 3 (Theorem 1 and Propositions 2 and 3) are exact analogs of their counterparts in Gerratana and Koçkesen (2012). However, the games to which these are applied are completely different. This becomes most transparent in applications (Section 6). Obtaining similarly sharp results in Gerratana and Koçkesen (2012) has been possible in a different class of games and the results are quite different. Furthermore, in the current paper we extend our results to (1) arbitrary extensive form games and mixed strategies; (2) to non-neutral third-parties; and (3) to the case of observable contracts.

Another related paper is Caillaud et al. (1995), which analyzes a game between two principal-agent hierarchies. In the first stage of their game each principal decides whether to publicly offer a contract to the agent; in the second stage each principal offers a secret contract to the agent, which, if accepted, overwrites the public contract that might have been offered in stage 1; in the third stage each agent receives payoff relevant information, decides whether to quit, and if he does not quit, he plays a normal form game with the other agent. Their main question is whether there exist equilibria of this game in which the principals choose not to offer a public contract in stage 1. If the answer to this question is no, then the interpretation is that contracts have commitment value. They show that contracts have commitment value if the market game stage is of Cournot type, but not if it is of

Bertrand type. The crucial difference between Caillaud et al. (1995) and our model is that they allow renegotiation only before the game begins, whereas in our setting renegotiation can happen both before and after the game begins.[6]

Finally, Bensaid and Gary-Bobo (1993) analyze a model in which the original game is a two-stage game and the initial contract can be renegotiated after player 1 chooses an action. However, in their model utility is not transferable between player 2 and the third-party. They show that, in a certain class of games, contracts with third parties have a commitment effect, even when they are renegotiable. We analyze a model with transferable utility and show that commitment effects still exist.

2 The Model

Our aim is to understand the effects of renegotiation-proof third-party contracts in extensive form games. In the main body of the paper we will do this in a particularly simple environment: two-stage games with private information, which we call the *original game*. The main reason we present our results for two-stage games is ease of exposition. Still, we should note that many models in economics such as the entry game, the Stackelberg game, and monopolistic screening are two-stage games, belong to this class of games. Furthermore, we show in Section 4.1 that our main result extends to arbitrary extensive form games with incomplete information as long as they satisfy an increasing differences property (see Definition 8).

We then allow one of the players to sign a contract with a third-party before the original game begins and call this new game the *game with third-party contracts*. The contract specifies a transfer between the player and the third-party as a function of the contractible outcomes of the original game. The crucial aspect of our model is the presence of asymmetric information between this player and the third-party during the renegotiation phase.

More precisely, we define the *original game*, denoted G, as follows: Nature chooses $\theta \in \Theta$ according to probability distribution $p \in \Delta(\Theta)$. After the move of Nature, player 1, without observing θ, chooses $a_1 \in A_1$. Lastly, player 2 observes (θ, a_1) and chooses $a_2 \in A_2$. We assume that A_1, A_2, and Θ are finite and let $p(\theta)$ denote the probability of Nature choosing θ. Payoff function of player $i \in \{1,2\}$ is given by $u_i : A \times \Theta \to \mathbb{R}$, where $A = A_1 \times A_2$.

The *game with third-party contracts* is a three player extensive form game described by the following sequence of events: Player 2 offers a contract $f : A \to \mathbb{R}$ to a third-party. The third-party accepts (denoted y) or rejects (denoted n) the contract. In case of rejection the game ends and the third-party receives a fixed payoff of $\delta \in \mathbb{R}$ while player 2 receives $-\infty$. In case of acceptance, player 1 and 2 play the original game. We assume that throughout the entire game θ remains the private information of player 2.

Since offering a contract that is rejected yields player 2 a very small payoff, the contract offer will be accepted in all equilibria. Therefore, we can omit the third-party's acceptance decision from histories and represent an outcome of the game with third-party contracts as (f, θ, a_1, a_2). The payoff functions in the game with contracts are given by $v_1\left(f, a_1, a_2, \theta\right) = u_1\left(a_1, a_2, \theta\right), v_2\left(f, a_1, a_2, \theta\right) = u_2\left(a_1, a_2, \theta\right) - f\left(a_1, a_2\right), v_3\left(f, a_1, a_2, \theta\right) = f\left(a_1, a_2\right)$, where v_3 is the payoff function of the third-party. Note that the payoff function of the third-party assumes that he is neutral towards the outcome of

[6] This is also related to the fact that they assume the agents play a simultaneous move game whereas we focus on sequential move games.

the game, i.e., he cares only about the transfer. In Section 4.2 we relax this assumption and allow the third-party also to have intrinsic preferences over the outcomes of the original game.

The game is with renegotiable contracts if the contracting parties can renegotiate the contract after player 1 plays a_1 and before player 2 chooses a_2. We assume that player 2, who is the informed party, initiates the renegotiation process by offering a new contract, which the third-party may accept or reject. If the third-party rejects the renegotiation offer g, then player 2 chooses $a_2 \in A_2$ and the outcome is payoff equivalent to (f, θ, a_1, a_2). If he accepts, then player 2 chooses $a_2 \in A_2$ and the outcome is payoff equivalent to (g, θ, a_1, a_2).

We say that the game is with observable contracts if the initial contract is observed by player 1. Otherwise, we say that the game is with unobservable contracts. In other words, there are four possible games with third-party contracts depending upon whether the contract is renegotiable or non-renegotiable and observable or unobservable. Given an original game G, we will denote the game with non-renegotiable and observable contracts with Γ_{NO}, non-renegotiable and unobservable contracts with Γ_{NU}, renegotiable and observable contracts with Γ_{RO}, and renegotiable and unobservable contracts with Γ_{RU}.

In any original game or game with contracts, a *behavior strategy* for player $i \in \{1,2,3\}$ is defined as a set of probability measures $\beta_i \equiv \{\beta_i(I) : I \in \mathscr{I}_i\}$, where $\mathscr{I}_i$ is the set of information sets of player i and $\beta_i(I)$ is defined on the set of actions available at information set I. One may write $\beta_i(h)$ for $\beta_i(I)$ for any history $h \in I$. By a *system of beliefs*, we mean a set $\mu \equiv \{\mu(I) : I \in \mathscr{I}_i$ for some $i\}$, where $\mu(I)$ is a probability measure on I. A pair (β, μ) is called an *assessment*. An assessment (β, μ) is said to be a *perfect Bayesian equilibrium* (PBE) if (1) each player's strategy is optimal at every information set given her beliefs and the other players' strategies; and (2) beliefs at every information set are consistent with observed histories and strategies.[7]

We will limit our analysis to pure behavior strategies, and hence a strategy profile of the original game G is given by $(b_1, b_2) \in A_1 \times A_2^{A_1 \times \Theta}$.[8] For any behavioral strategy profile (b_1, b_2) of G, define the expected payoff of player $i = 1,2$ as $U_i(b_1, b_2) = \sum_{\theta \in \Theta} p(\theta) u_i(b_1, b_2(b_1, \theta), \theta)$ and the best response correspondences as $BR_1(b_2) = \text{argmax}_{a_1 \in A_1} U_1(a_1, b_2)$ for all $b_2 \in A_2^{A_1 \times \Theta}$ and $BR_2(a_1, \theta) = \text{argmax}_{a_2 \in A_2} u_2(a_1, a_2, \theta)$ for all $(a_1, \theta) \in A_1 \times \Theta$. We say that a strategy profile (b_1^*, b_2^*) is a *Bayesian Nash equilibrium* of G if $b_1^* \in BR_1(b_2^*)$ and $b_2^*(b_1^*, \theta) \in BR_2(b_1^*, \theta)$ for all θ. The difference between a perfect Bayesian equilibrium and a Bayesian Nash equilibrium, of course, is that the former requires player 2 to best respond to every action of player 1, whereas the latter requires best response to only the equilibrium action. Therefore, every perfect Bayesian equilibrium is a Bayesian Nash equilibrium but not conversely.

For any behavior strategy profile (b_1, b_2) in G, we say that an assessment (β, μ) in $\Gamma_k(G)$, $k = NO, NU, RO, RU$, *induces* (b_1, b_2) if in $\Gamma_k(G)$ player 1 plays according to b_1 and, after the equilibrium contract, player 2 plays according to b_2.[9]

Our ultimate aim is to characterize renegotiation-proof equilibria, in which the equilibrium contract is not renegotiated after any history.[10] More precisely,

[7]See Fudenberg and Tirole (1991) for a precise definition of perfect Bayesian equilibrium.

[8]In Section 4.1 we relax this and allow also mixed strategies. This introduces some technical difficulties but our main results go through.

[9]Note that in $\Gamma_{RO}(G)$ and $\Gamma_{RU}(G)$, player 2 may choose an action $a_2 \in A_2$ either without renegotiating the initial contract or after attempting renegotiation.

[10]We follow the previous literature in our definition of renegotiation-proof equilibrium. See, for example, Maskin and Tirole (1992) and Beaudry and Poitevin (1995).

Definition 1 (Renegotiation-Proof Equilibrium)**.** A perfect Bayesian equilibrium (β^*, μ^*) of $\Gamma_{RO}(G)$ and $\Gamma_{RU}(G)$ is *renegotiation-proof* if the equilibrium contract is not renegotiated after any $a_1 \in A_1$ and $\theta \in \Theta$.

We say that a strategy profile (b_1, b_2) of the original game G can be *supported with observable and non-renegotiable contracts* if there exists a perfect Bayesian equilibrium of $\Gamma_{NO}(G)$ that induces (b_1, b_2). Similarly, a strategy profile (b_1, b_2) of the original game G can be *supported with observable renegotiation-proof contracts* if there exists a renegotiation-proof perfect Bayesian equilibrium of $\Gamma_{RO}(G)$ that induces (b_1, b_2). Similarly for unobservable and non-renegotiable and unobservable renegotiation-proof contracts.

AN EXAMPLE: ENTRY DETERRENCE

In order to illustrate our main query as well as some of our results later on, we introduce a very simple entry game in this section (See Figure 1). Player 1 is a potential entrant, who may enter (E) or stay out (O) and player 2, who is the incumbent, may fight (F) or accommodate (A) entry.

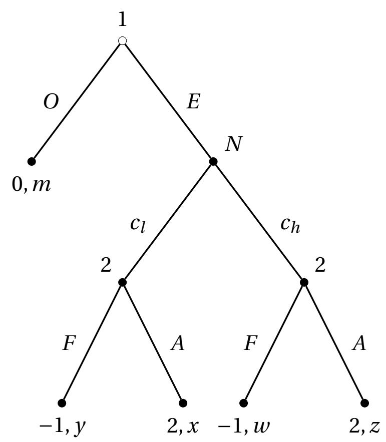

Figure 1: Entry Game

We assume that fighting is costly, and it is costlier for the high cost incumbent (type c_h) than for the low cost (type c_l): $z - w > x - y > 0$. The entrant believes that the incumbent's type is low cost with probability $p \in (0,1)$.

The unique perfect Bayesian equilibrium (PBE) of this game is (E, AA), i.e., the entrant enters and both types of the incumbent accommodate. We assume that the monopoly profit is larger than the highest possible profit following entry, i.e., $m > z$. In other words, the incumbent would benefit from deterring entry, and one way of achieving this would be to sign a contract with a third-party that makes fighting optimal. For example, the following contract makes playing FF optimal: $f(F) = \delta, f(A) = \delta + (z - w)$. Is such a contract renegotiation-proof? If not, can entry still be deterred with renegotiation-proof contracts? In what follows we will answer these questions and also characterize the equilibrium outcomes that can be supported with third-party contracts under different assumptions regarding their observability and renegotiation-proofness.

3 Renegotiation-Proof Contracts

In this section we will provide results that help identify the set of outcomes of any original game G that can be supported by renegotiation-proof perfect Bayesian equilibria of the game with observable (or unobservable) and renegotiable contracts.

In order to decide whether to accept a new contract offer in the renegotiation phase of the game with renegotiable contracts, the third-party forms beliefs regarding player 2's strategy under the new contract and compares his payoffs from the old and the new contracts. In equilibrium, these beliefs must be such that player 2's strategies are sequentially rational, i.e., incentive compatible, under the new contract. Let $\mathscr{C} = \mathbb{R}^{A_1 \times A_2}$ and define incentive compatibility as a property of any contract-strategy pair $(f, b_2) \in \mathscr{C} \times A_2^{A_1 \times \Theta}$.

Definition 2 (Incentive Compatibility)**.** $(f, b_2) \in \mathscr{C} \times A_2^{A_1 \times \Theta}$ is *incentive compatible* if

$$u_2(a_1, b_2(a_1,\theta),\theta) - f(a_1, b_2(a_1,\theta)) \geq u_2(a_1, b_2(a_1,\theta'),\theta) - f(a_1, b_2(a_1,\theta')) \text{ for all } a_1 \in A_1 \text{ and } \theta,\theta' \in \Theta.$$

We say that a strategy b_2 is incentive compatible if there is a contract f such that (f, b_2) is incentive compatible. We can obtain a sharp characterization of incentive compatible strategies if we impose more structure on the original game. To this end, let $\succsim_\theta$ be a linear order on Θ and $\succsim_2$ a linear order on A_2, and denote their asymmetric parts by $\succ_\theta$ and $\succ_2$, respectively.

Definition 3 (Increasing Differences)**.** $u_2 : A_1 \times A_2 \times \Theta \to \mathbb{R}$ is said to have *increasing differences* in $(\succsim_\theta, \succsim_2)$ if for all $a_1 \in A_1$, $\theta \succsim_\theta \theta'$ and $a_2 \succsim_2 a_2'$ imply that $u_2(a_1, a_2,\theta) - u_2(a_1, a_2,\theta') \geq u_2(a_1, a_2',\theta) - u_2(a_1, a_2',\theta')$. It is said to have *strictly increasing differences* if $\theta \succ_\theta \theta'$ and $a_2 \succ_2 a_2'$ imply that $u_2(a_1, a_2,\theta) - u_2(a_1, a_2,\theta') > u_2(a_1, a_2',\theta) - u_2(a_1, a_2',\theta')$.

Definition 4 (Increasing Strategies)**.** $b_2 : A_1 \times \Theta \to A_2$ is called *increasing* in $(\succsim_\theta, \succsim_2)$ if for all $a_1 \in A_1$, $\theta \succsim_\theta \theta'$ implies that $b_2(a_1,\theta) \succsim_2 b_2(a_1,\theta')$. Denote the set of all increasing b_2 by B_2^+.

For the rest of the paper, we restrict attention to games G in which there exist a linear order on Θ and a linear order on A_2 such that u_2 has strictly increasing differences in $(\succsim_\theta, \succsim_2)$. Standard arguments show that under increasing differences, incentive compatibility implies that b_2 is increasing. The following proposition states this result and shows that its converse also holds.

Proposition 1. *If $u_2 : A_1 \times A_2 \times \Theta \to \mathbb{R}$ has strictly increasing differences, then a strategy $b_2 : A_1 \times \Theta \to A_2$ is incentive compatible if and only if it is increasing.*

The only if part follows from a standard argument in contract theory. In order to prove the if part fix an arbitrary $a_1 \in A_1$, let the number of elements of Θ be n, and order its elements so that $\theta^n \succsim_\theta \theta^{n-1} \succsim_\theta \cdots \theta^2 \succsim_\theta \theta^1$. For any contract-strategy pair (f, b_2), define $f(a_1)_j = f(a_1, b_2(a_1,\theta^j))$, $j = 1,\ldots,n$, and let, with an abuse of notation, $f(a_1) \in \mathbb{R}^n$ be the vector whose jth component is given by $f(a_1)_j$. When u_2 has increasing differences, incentive compatibility of (f, b_2) is equivalent to the local upward and downward constraints:[11]

$$\begin{aligned} f(a_1)_j - f(a_1)_{j+1} &\leq u_2(a_1, b_2(a_1,\theta^j),\theta^j) - u_2(a_1, b_2(a_1,\theta^{j+1}),\theta^j), \quad j = 1,\ldots,n-1 \\ -f(a_1)_{j-1} + f(a_1)_j &\leq u_2(a_1, b_2(a_1,\theta^j),\theta^j) - u_2(a_1, b_2(a_1,\theta^{j-1}),\theta^j), \quad j = 2,\ldots,n \end{aligned}$$

[11] See, for example, Bolton and Dewatripont (2005), p. 78.

For any $a_1 \in A_1$, we can write these inequalities in matrix form as $Df(a_1) \leq \vec{U}_2(a_1, b_2)$, where D is a matrix of coefficients and $\vec{U}_2(a_1, b_2)$ a column vector with $2(n-1)$ components, whose component $2j-1$ is given by

$$\vec{U}_2(a_1, b_2)_{2j-1} = u_2(a_1, b_2(a_1, \theta^j), \theta^j) - u_2(a_1, b_2(a_1, \theta^{j+1}), \theta^j)$$

and component $2j$ is given by

$$\vec{U}_2(a_1, b_2)_{2j} = u_2(a_1, b_2(a_1, \theta^{j+1}), \theta^{j+1}) - u_2(a_1, b_2(a_1, \theta^j), \theta^{j+1})$$

Therefore, the proof will be completed if we can show that if u_2 has strictly increasing differences and b_2 increasing, then there exists $f(a_1) \in \mathbb{R}^n$ such that $Df(a_1) \leq \vec{U}_2(a_1, b_2)$. This follows easily from Gale's theorem for linear inequalities (Mangasarian (1994), p. 33).

We next define our renegotiation-proofness concept, which follows from the definition of renegotiation-proof perfect Bayesian equilibrium (Definition 1).

Definition 5 (Renegotiation-Proofness)**.** We say that $(f, b_2^*) \in \mathscr{C} \times A_2^{A_1 \times \Theta}$ is *renegotiation-proof* if for all $a_1 \in A_1$ and $\theta \in \Theta$ for which there exists an incentive compatible (g, b_2) such that

$$u_2(a_1, b_2(a_1, \theta), \theta) - g(a_1, b_2(a_1, \theta)) > u_2(a_1, b_2^*(a_1, \theta), \theta) - f(b_2^*(a_1, \theta)) \tag{1}$$

there exists a $\theta' \in \Theta$ such that

$$f(a_1, b_2^*(a_1, \theta')) \geq g(a_1, b_2(a_1, \theta')) \tag{2}$$

In words, if, for some (θ, a_1), there is a contract g and an incentive compatible continuation play b_2 such that player 2 prefers g over f (i.e., (1) holds), there must exist a belief of the third-party (over θ) under which it is optimal to reject g, which is implied by (2).[12]

Finally, we define a renegotiation-proof strategy as,

Definition 6 (Renegotiation-Proof Strategy)**.** A strategy $b_2 \in A_2^{A_1 \times \Theta}$ is *renegotiation-proof* if there exists an $f \in \mathscr{C}$ such that (f, b_2) is incentive compatible and renegotiation-proof. Denote the set of all renegotiation-proof strategies by B_2^R

It is not difficult to see that Definitions 5 and 6 are indeed the correct definitions to work with, in the sense that they identify the conditions that any contract f and strategy b_2 must satisfy to be part of a renegotiation-proof perfect Bayesian equilibrium of $\Gamma_{RO}(G)$ or $\Gamma_{RU}(G)$. Indeed, if a strategy b_2 of the original game is not renegotiation-proof, then there is no renegotiation-proof perfect Bayesian equilibrium (of the game with renegotiable contracts) in which a contract f is offered and b_2 is played without renegotiating f. This simply follows from the fact that if (f, b_2) is not renegotiation-proof, then there is (a_1, θ) and a contract g that would be accepted for any belief of the third-party at the renegotiation stage and increase player 2's payoff. In other words, f will be renegotiated after (a_1, θ) and therefore the equilibrium is not renegotiation-proof. In fact, the converse of that statement also holds: If b_2 is renegotiation-proof, we can construct a perfect Bayesian equilibrium of the game with renegotiable contracts in which the equilibrium contract is not renegotiated after any a_1 and θ. Of

[12] This definition allows beliefs to be arbitrary following an off-the-equilibrium renegotiation offer. An alternative definition would be to require the beliefs to satisfy intuitive criterion. In Section 4.3 we show that our results go through with minor modifications when we adopt this stronger version.

course, the equilibrium contract and b_2 will also have to satisfy other conditions for them to be part of an equilibrium, but these would depend on whether the contracts are observable or unobservable, an issue which we will address in Section 5.

Our main result in this section characterizes renegotiation-proof contracts and strategies. In order to understand this result one should first realize that condition (2) in Definition 5 is satisfied trivially if the strategy b_2 does not lead to a higher surplus for the contracting parties after (a_1, θ). In other words, for each a_1 and $i = 1, \dots n$, we need to check renegotiation-proofness of (f, b_2^*) only against strategies that belong to the following set:

$$\mathfrak{B}(a_1, i, b_2^*) = \{b_2 \in A_2^{A_1 \times \Theta} : b_2 \text{ is increasing and } u_2(a_1, b_2(a_1, \theta^i), \theta^i) > u_2(a_1, b_2^*(a_1, \theta^i), \theta^i)\}. \quad (3)$$

Third, by Definition 5, (f, b_2^*) is *not* renegotiation-proof if and only if there exist $a_1 \in A_1$, $i = 1, \dots n$, and incentive compatible (g, b_2) such that $u_2(a_1, b_2(a_1, \theta^i), \theta^i) - g(a_1)_i > u_2(a_1, b_2^*(a_1, \theta_i), \theta_i) - f(a_1)_i$ and $g(a_1)_j > f(a_1)_j$ for all $j = 1, \dots, n$. As we have discussed after Proposition 1, when u_2 has increasing differences, incentive compatibility of (g, b_2) is equivalent to $Dg(a_1) \le \vec{U}_2(a_1, b_2)$. Therefore, (f, b_2^*) is not renegotiation-proof if and only if there exist a_1, i, b_2 and $\varepsilon \in \mathbb{R}^n$ such that

$$D(f(a_1) + \varepsilon) \le \vec{U}_2(a_1, b_2), \quad \varepsilon_i < u_2(a_1, b_2(a_1, \theta^i), \theta^i) - u_2(a_1, b_2^*(a_1, \theta^i), \theta^i), \quad \varepsilon \gg 0$$

These conditions can be written as $[Ax \gg 0, Cx \ge 0$ has a solution $x]$, once the vector x and matrices A and C are appropriately defined. Motzkin's theorem of the alternative then implies that the necessary and sufficient condition for being renegotiation-proof is $[A'y_1 + C'y_2 = 0, y_1 > 0, y_2 \ge 0$ has a solution $y_1, y_2]$. The fact that u_2 has increasing differences can then be used to prove the equivalence of this condition to the one stated in the following theorem.

Theorem 1. *$(f, b_2^*) \in \mathscr{C} \times A_2^{A_1 \times \Theta}$ is renegotiation-proof if and only if for any $a_1 \in A_1$, $i \in \{1, 2, \dots, n\}$, and $b_2 \in \mathfrak{B}(a_1, i, b_2^*)$ there exists a $k \in \{1, 2, \dots, i-1\}$ such that*

$$u_2(a_1, b_2(a_1, \theta^i), \theta^i) - u_2(a_1, b_2^*(a_1, \theta^i), \theta^i) + \sum_{j=k}^{i-1} \vec{U}_2(a_1, b_2)_{2j-1} \le f(a_1)_k - f(a_1)_i \quad (4)$$

or there exists an $l \in \{i+1, i+2, \dots, n\}$ such that

$$u_2(a_1, b_2(a_1, \theta^i), \theta^i) - u_2(a_1, b_2^*(a_1, \theta^i), \theta^i) + \sum_{j=i+1}^{l} \vec{U}_2(a_1, b_2)_{2(j-1)} \le f(a_1)_l - f(a_1)_i \quad (5)$$

Proof. Similar to Theorem 2 in Gerratana and Koçkesen (2012). □

Theorem 1 characterizes the conditions for which $\left(f, b_2^*\right)$ is renegotiation-proof. Our next step is to find conditions for a strategy b_2^* to be supported with renegotiation-proof contracts. The following definition facilitates the exposition.

Definition 7. For any a_1, $i = 1, \dots, n$ and $b_2 \in \mathfrak{B}(a_1, i, b_2^*)$ we say that $m(b_2) \in \{1, 2, \dots, n\}$ is a *blocking type* if

$$u_2(a_1, b_2(a_1, \theta^i), \theta^i) - u_2(a_1, b_2^*(a_1, \theta^i), \theta^i) \le \sum_{j=m(b_2)}^{i-1} \left[\vec{U}_2(a_1, b_2^*)_{2j-1} - \vec{U}_2(a_1, b_2)_{2j-1}\right] \quad (6)$$

or

$$u_2(a_1, b_2(a_1, \theta^i), \theta^i) - u_2(a_1, b_2^*(a_1, \theta^i), \theta^i) \leq \sum_{j=i+1}^{m(b_2)} \left[\vec{U}_2(a_1, b_2^*)_{2(j-1)} - \vec{U}_2(a_1, b_2)_{2(j-1)}\right] \quad (7)$$

We obtain the following necessary conditions for a strategy b_2^* to be renegotiation-proof.

Proposition 2. *A strategy $b_2^* \in A_2^{A_1 \times \Theta}$ is renegotiation-proof only if for any $a_1 \in A_1$, $i \in \{1, 2, \dots, n\}$, and $b_2 \in \mathfrak{B}(a_1, i, b_2^*)$ there is a blocking type.*

Proof. Similar to Proposition 3 in Gerratana and Koçkesen (2012). □

The above condition becomes also sufficient for renegotiation-proofness with an additional requirement about the relation of blocking types for different renegotiation opportunities.

Proposition 3. *A strategy $b_2^* \in A_2^{A_1 \times \Theta}$ is renegotiation-proof if for any $a_1 \in A_1$, $i \in \{1, 2, \dots, n\}$, and $b_2 \in \mathfrak{B}(a_1, i, b_2^*)$ there is a blocking type $m(b_2^i)$ such that $k < l$, $m(b_2^k) > k$, and $m(b_2^l) < l$ imply $m(b_2^k) \leq m(b_2^l)$.*

Proof. Similar to Proposition 4 in Gerratana and Koçkesen (2012). □

The conditions given in Proposition (2) and (3) coincide when player 2 has only two types. Therefore, Proposition (2) is a full characterization result for such games. Although, they fall short of providing a full characterization in games with more than two types, they help us do so in environments with more structure as we demonstrate in Section 6.

EXAMPLE: ENTRY DETERRENCE

Let $c_h \succ_\theta c_l$ and $A \succ_2 F$ and observe that $z - w > x - y$ implies that u_2 has strictly increasing differences. Proposition 1 therefore implies that the set of incentive compatible strategies are $\{FF, FA, AA\}$. Are there strategies renegotiation-proof? AA is clearly renegotiation-proof because both types are best responding and hence $\mathfrak{B}(E, c_l, AA) = \mathfrak{B}(E, c_h, AA) = \emptyset$. How about FF? For both types playing A is a better response and hence $\mathfrak{B}(E, c_l, FF) = \{AA\}$ and $\mathfrak{B}(E, c_h, FF) = \{FA, AA\}$. Is there a blocking type for c_l, i.e., does (7) hold for $m(b_2) = c_h$? Since

$$u_2(E, A, c_l) - u_2(E, F, c_l) = x - y > u_2(E, F, c_h) - u_2(E, F, c_h) - (u_2(E, A, c_h) - u_2(E, A, c_h)) = 0$$

the answer is no, i.e., FF is not renegotiation-proof. Is FA renegotiation-proof? In this case $\mathfrak{B}(E, c_l, FA) = \{AA\}$ and $\mathfrak{B}(E, c_h, FA) = \emptyset$. Is there a blocking type for c_l? Since

$$u_2(E, A, c_l) - u_2(E, F, c_l) = x - y \leq u_2(E, A, c_h) - u_2(E, F, c_h) - (u_2(E, A, c_h) - u_2(E, A, c_h)) = z - w$$

the answer is yes. Therefore, the set of renegotiation-proof strategies is $\{FA, AA\}$. In other words, renegotiation-proofness in this example is satisfied whenever the high cost type best responds. Also note that for the high cost type, not best responding is costlier, i.e., $z - w > x - y$. Credible commitment, in this example, requires best responding when it is very costly no to do so. Finally, an example of a contract that supports FA is $f(F) = \delta$, $f(A) = \delta + (x - y)$.

4 Extensions

So far we have assumed that the original game has only two stages and conducted the equilibrium analysis in pure strategies. Furthermore, we have assumed that the third-party is neutral. In this section we show that all our main results can be generalized to a much more general class of extensive form games with incomplete information and they are true in mixed strategy equilibria as well. We also discuss how our results are modified when third-parties are not neutral. Finally, we present the implications of a stronger definition of renegotiation-proofness.

4.1 General Extensive Form Games

Although we have stated our results for two stage games in which only player 2 has private information and has the right to sign a contract with a third-party, we can generalize them to arbitrary finite extensive form games with incomplete information and perfect recall. The only restriction we impose is that players' payoff functions in the original game exhibit increasing differences in a sense that we will make precise.

Define the original game G as an extensive form game with incomplete information in which player $i \in \{1,\dots,n\}$ privately learns his type $\theta_i \in \Theta_i$ at the beginning of the game. Assume that Nature chooses types independently and let $|\Theta_i| = n_i$. After the types are determined players start taking actions. We denote the set of histories (excluding the moves of Nature at the beginning) by H and denote a typical history by $h = (a_0, a_1, a_2, \dots, a_k)$, where a_0 is the initial node (or empty history) and a_j denotes the jth action taken in this history. Payoff function of player i is given by $u_i : Z \times \Theta \to \mathbb{R}$, where Z is the set of (finite) terminal histories and Θ is the set of all type profiles $\theta = (\theta_1, \dots, \theta_n)$.

The set of pure strategies of player i is given by S_i and a mixed strategy for player i is a probability distribution over S_i for each θ_i, i.e., a mapping $\sigma_i : \Theta_i \to \Delta(S_i)$. A pure strategy profile is denoted by s and mixed strategy profile by σ. Denote the set of all information sets at which player i moves by $\mathscr{I}_i$ and the set of all information sets in the game by $\mathscr{I}$. At any information set $I \in \mathscr{I}_i$, player i has a set of pure strategies available for the rest of the game, denoted $S_i|_I$ and defined as the restriction of S_i to information sets of player i that follows (and includes) I. A belief system is a collection of probability measures $m = \{m(I) : I \in \mathscr{I}\}$, where $m(I)$ for $I \in \mathscr{I}_i$ is defined over $I \times \Theta_{-i}$. A pair (σ, m) is called an assessment.[13] We say that an assessment is consistent if beliefs at every information set are derived from prior beliefs, observed histories, and strategies using Bayes' Law whenever possible.

We assume that for every player i and information set $I \in \mathscr{I}_i$, $S_i|_I$ is a chain. In other words, there is a binary relation $\succsim_i$ on Θ_i and $\succsim_{si}$ on S_i that is reflexive, antisymmetric, transitive, and complete.[14] We denote the asymmetric parts of $\succsim_i$ and $\succsim_{si}$ by $\succ_i$ and $\succ_{si}$, respectively.

Fix a player i and an information set $I \in \mathscr{I}_i$. Given an history $h \in I$, if the type profile is θ, player i plays $s_i \in S_i|_I$, and other players play $s_{-i} \in S_{-i}|_I$, payoff of player i can be written as $u_i(h, s_i, s_{-i}, \theta_i, \theta_{-i})$.[15]

Definition 8 (Increasing Differences)**.** We say that an original game G has increasing differences if for

[13]The original game as well as the games with contracts that we will define shortly have perfect recall and therefore for every behavior strategy there is an outcome equivalent mixed strategy and vice versa. Therefore, we are free to work with either the behavior or mixed strategies, whichever more convenient.

[14]Note that any chain is a lattice. Also, we omit the dependence of $\succsim_{si}$ on I to avoid notational clutter.

[15]We can do that by finding the terminal history $z(h, s_i, s_{-i})$ that is reached when players play according to (s_i, s_{-i}) after h and defining $u_i(h, s_i, s_{-i}, \theta_i, \theta_{-i}) = u_i(z(h, s_i, s_{-i}), \theta_i, \theta_{-i})$.

any $i \in N$, $I \in \mathscr{I}_i$, $\theta_i \succsim_i \theta'_i$ and $s_i \succsim_{si} s'_i$ imply that

$$u_i(h, s_i, s_{-i}, \theta_i, \theta_{-i}) - u_i(h, s'_i, s_{-i}, \theta_i, \theta_{-i}) \geq u_i(h, s_i, s_{-i}, \theta'_i, \theta_{-i}) - u_i(h, s'_i, s_{-i}, \theta'_i, \theta_{-i})$$

for all $h \in I$, $s_{-i} \in S_{-i}|_I$, and $\theta_{-i} \in \Theta_{-i}$. It is said to have *strictly increasing differences* if $\theta_i \succ_i \theta'_i$ and $s_i \succ_{si} s'_i$ imply that

$$u_i(h, s_i, s_{-i}, \theta_i, \theta_{-i}) - u_i(h, s'_i, s_{-i}, \theta_i, \theta_{-i}) > u_i(h, s_i, s_{-i}, \theta'_i, \theta_{-i}) - u_i(h, s'_i, s_{-i}, \theta'_i, \theta_{-i})$$

for all $h \in I$, $s_{-i} \in S_{-i}|_I$, and $\theta_{-i} \in \Theta_{-i}$.

In other words, an original game G has increasing differences if the payoff functions have increasing differences in $(s_i, \theta_i) \in S_i|_I \times \Theta_i$ at every information set I irrespective for how the other players play and what the types of the other players are. We assume that the original game G has strictly increasing differences. Examples of games with increasing differences include repeated ultimatum bargaining and chain store games.

The induced game with non-renegotiable and renegotiable contracts are straightforward generalizations of their counterparts for two stage games. Each player i independently offers a contract $f_i : Z \to \mathbb{R}$ to a distinct third-party t_i, who accepts or rejects it. In case of rejection the game ends, t_i receives a fixed payoff of $\delta_i \in \mathbb{R}$, and player i receives $-\infty$. In case of acceptance Nature chooses θ and players in N play G.

We assume that only player i observes his type θ_i, and that this is the only source of asymmetric information between i and the third-party t_i. In other words, at any point in the game both the main player and his third-party observe the same histories. The payoff functions are given by $v_i(f, z, \theta = u_i(z, \theta) - f_i(z)$, $v_{t_i}(f, z, \theta) = f_i(z)$.

One of the conditions that strategies must satisfy in any perfect Bayesian equilibrium of the game with contracts, is that player i's strategy must be sequentially rational, or incentive compatible, under the contract. Increasing differences imply that this is equivalent to strategies being increasing. Define a mixed strategy as increasing if any pure strategy in its support is increasing in the type. We have the following counterpart to Proposition 1.

Proposition 4. *If the original game G has strictly increasing differences, then a mixed strategy is incentive compatible if and only if it is increasing.*

Proof. Omitted. □

The game is with renegotiable contracts if the contracting parties can renegotiate the contract at any point throughout the game. At any information set $I \in \mathscr{I}_i$ player i either offers a new contract $g_i : Z \to \mathbb{R}$ to the third-party t_i or chooses an action. If player i offers a new contract, the third-party either accepts or rejects it.

We retain the same definition of renegotiation-proofness. Fix a consistent assessment (σ, m). We say that a contract strategy pair (f_i, σ_i) is renegotiation-proof at (σ_{-i}, m), if whenever there is a contract g_i and an incentive compatible continuation play σ'_i such that player i of type θ_i strictly prefers (g_i, σ'_i) over (f_i, σ_i) at information set I, there must exist a type θ'_i for which expected transfers under (f_i, σ_i) at least as high as the transfers under (g_i, σ'_i).

In order to state the counterpart to our main result for renegotiation-proof contracts we need a few more definitions. Fix a consistent assessment (σ,m) and let $U_i^{\sigma,m}(\sigma_i',\theta_i|I)$ be the expected payoff of player i of type θ_i to playing mixed strategy $\sigma_i' \in \Delta(S_i|_I)$ conditional on reaching information set I. Similarly, let $F_i^{\sigma,m}(\sigma_i'|I)$ be the expected transfers.

For any $i \in N$, $I \in \mathscr{I}_i$, consistent assessment (σ,m), and $\sigma_i' : \Theta_i \to \Delta(S_i|_I)$ define $\vec{U}_i^{\sigma,m}(\sigma_i'|I)$ as a column vector with $2(n_i-1)$ components, where component $2j-1$ is given by $U_i^{\sigma,m}(\sigma_i'(\theta_i^j),\theta^j|I) - U_i^{\sigma,m}(\sigma_i'(\theta_i^{j+1}),\theta^j|I)$ and component $2j$ is given by $U_i^{\sigma,m}(\sigma_i'(\theta_i^{j+1}),\theta^{j+1},|I) - U_i^{\sigma,m}(\sigma_i'(\theta_i^j),\theta^{j+1}|I)$, $j = 1,2,\ldots,n_i-1$. Similarly, define $\vec{F}_i^{\sigma,m}(\sigma_i'|I)$ as the n_i vector whose jth component is given by $F_i^{\sigma,m}(\sigma_i'(\theta_i^j)|I)$, $j = 1,2,\ldots,n_i-1$.[16] Let

$$\widehat{\Sigma}_i(I,j,\sigma_i) = \{\sigma_i' : \Theta_i \to \Sigma_i|_I : \sigma_i' \text{ is increasing and } U_i^{\sigma,m}(\sigma_i'(\theta_i^j),\theta_i^j|I) > U_i^{\sigma,m}(\sigma_i(\theta_i^j),\theta_i^j|I)\}$$

be the set of all mixed strategies at information set I that is increasing and increases the payoff of player i type j over his payoff under σ_i.

Theorem 1 generalizes in a quite straightforward way:

Theorem 2. *Fix a consistent assessment (σ,m) and $i \in N$. $(f_i,\sigma_i) \in \mathscr{C} \times \Sigma_i$ is* renegotiation-proof *at (σ_{-i},m) if and only if for any $I \in \mathscr{I}_i$, $j = 1,\ldots,n_i$, and $\sigma_i' \in \widehat{\Sigma}_i(I,j,\sigma_i)$ there exists a $k \in \{1,2,\ldots,j-1\}$ such that*

$$U_i^{\sigma,m}(\sigma_i'(\theta_i^j),\theta_i^j|I) - U_i^{\sigma,m}(\sigma_i(\theta_i^j),\theta_i^j|I) + \sum_{t=k}^{j-1} \vec{U}_i^{\sigma,m}(\sigma_i'|I)_{2t-1} \le \vec{F}_i^{\sigma,m}(\sigma_i|I)_k - \vec{F}_i^{\sigma,m}(\sigma_i|I)_j \tag{8}$$

or there exists an $l \in \{j+1, j+2,\ldots,n_i\}$ such that

$$U_i^{\sigma,m}(\sigma_i'(\theta_i^j),\theta_i^j|I) - U_i^{\sigma,m}(\sigma_i(\theta_i^j),\theta_i^j|I) + \sum_{t=j+1}^{l} \vec{U}_i^{\sigma,m}(\sigma_i'|I)_{2(t-1)} \le \vec{F}_i^{\sigma,m}(\sigma_i|I)_l - \vec{F}_i^{\sigma,m}(\sigma_i|I)_j \tag{9}$$

Proof. Omitted. □

Once the definition of a blocking type is appropriately modified (see Definition 7), Propositions 2 and 3 also generalize in a straightforward manner.

4.2 Interested Third-Party

In our model we assumed that the third-party has no interest in the outcome of the original game other than the transfer from (or to) player 2. This is obviously not always the case with third-party contracts. For example, the government in its contractual relationship with a central bank is definitely interested in the outcome of the game between the central bank and the public. Similarly, the European Union in its contractual relationships with Airbus is interested in the entry game played by Airbus and Boeing. We can easily think of many other instances of games with third-party contracts in which the third-party itself is interested in the outcome of the game. How do our results change if this is the case? The answer turns out to be straightforward and intuitive.

Let $u_3(a_1,a_2,\theta)$ be the third-party's payoff function so that under contract f his payoff would be $u_3(a_1,a_2,\theta) + f(a_1,a_2)$. We say that (f,b_2^*) is renegotiation-proof if for all $a_1 \in A_1$ and $\theta \in \Theta$ for which

[16] Note that independence of types across players implies that the beliefs of player i over h and θ_{-i}, i.e., $m(h,\theta_{-i}|I)$, do not depend on θ_i. For the same reason expected transfers to pure strategy s_i do not depend on θ_i.

there exists an incentive compatible (g, b_2) such that

$$u_2(a_1, b_2(a_1,\theta),\theta) - g(a_1, b_2(a_1,\theta)) > u_2(a_1, b_2^*(a_1,\theta),\theta) - f(b_2^*(a_1,\theta))$$

there exists a $\theta' \in \Theta$ such that

$$u_3(a_1, b_2^*(a_1,\theta'),\theta') + f(a_1, b_2^*(a_1,\theta')) \geq u_3(a_1, b_2(a_1,\theta'),\theta') + g(a_1, b_2(a_1,\theta'))$$

In the model with neutral third-party a renegotiation opportunity arises whenever there is an increasing strategy that increases player 2's payoff $u_2(a_1, a_2,\theta)$, which is the total surplus available to player 2 and the third-party in that model. When the third-party is no longer neutral, total surplus available becomes $u_2(a_1, a_2,\theta) + u_3(a_1, a_2,\theta)$. Accordingly, a renegotiation opportunity arises whenever there is an increasing strategy that increases total surplus $u_2(a_1, a_2,\theta) + u_3(a_1, a_2,\theta)$. Therefore, we modify the definition of $\mathfrak{B}(a_1, i, b_2^*)$ as the set of strategies b_2 that are increasing and satisfy

$$u_2(a_1, b_2(a_1,\theta^i),\theta^i) + u_3(a_1, b_2(a_1,\theta^i),\theta^i) > u_2(a_1, b_2^*(a_1,\theta^i),\theta^i) + u_3(a_1, b_2^*(a_1,\theta^i),\theta^i)\}.$$

We can now state the modified version of Theorem 1:

Theorem 3. *(f, b_2^*) is renegotiation-proof if and only if for any $a_1 \in A_1$, $i \in \{1,2,\ldots,n\}$, and $b_2 \in \mathfrak{B}(a_1, i, b_2^*)$ there exists a $k \in \{1,2,\ldots,i-1\}$ such that*

$$u_2(a_1, b_2(a_1,\theta^i),\theta^i) - u_2(a_1, b_2^*(a_1,\theta^i),\theta^i) + u_3(a_1, b_2(a_1,\theta^k),\theta^k) - u_3(a_1, b_2^*(a_1,\theta^k),\theta^k) + \sum_{j=k}^{i-1} \vec{U}_2(a_1, b_2)_{2j-1} \leq f(a_1)_k - f(a_1)_i \quad (10)$$

or there exists an $l \in \{i+1, i+2,\ldots,n\}$ such that

$$u_2(a_1, b_2(a_1,\theta^i),\theta^i) - u_2(a_1, b_2^*(a_1,\theta^i),\theta^i) + u_3(a_1, b_2(a_1,\theta^l),\theta^l) - u_3(a_1, b_2^*(a_1,\theta^l),\theta^l) \sum_{j=i+1}^{l} \vec{U}_2(a_1, b_2)_{2(j-1)} \leq f(a_1)_l - f(a_1)_i \quad (11)$$

Proof. Omitted. ◻

Note that an interested third-party introduces two changes into the result: First, a renegotiation opportunity arises only if it increases the total surplus rather than just player 2's payoff. This might in fact help a contract become renegotiation-proof, if, for example, the third-party and player 2 have completely opposite preferences. Second, compared with (4) and (5), inequalities (10) and (11) have extra terms on the left hand side, which might help or hurt a contract become renegotiation-proof depending upon the sign of those terms.

Again, once the definition of a blocking type is appropriately modified, Propositions 2 and 3 can be easily generalized to the case of non-neutral third-party.

4.3 Strong Renegotiation-Proofness

Our definition of renegotiation-proofness follows directly from the assumed game form for the renegotiation procedure, i.e., player 2, who is the the informed party, makes a new contract offer and the third-party, who is uninformed, accepts or rejects. In a renegotiation-proof equilibrium, the contract is never renegotiated, and therefore any renegotiation offer is an out-of-equilibrium event. This allows us to specify the beliefs of the third-party freely after a new contract offer. This may be found unreasonable and a more plausible alternative could be to require beliefs satisfy the conditions specified in the intuitive criterion as defined by Cho and Kreps (1987).

In our setting, intuitive criterion requires that, given an equilibrium contract strategy pair (f, b_2^*) and following a renegotiation offer (g, b_2), beliefs put positive probability only on types for which (g, b_2) is not equilibrium-dominated, i.e., only on those types θ' for which

$$u_2(a_1, b_2(a_1,\theta'),\theta') - g(a_1, b_2(a_1,\theta')) \geq u_2(a_1, b_2^*(a_1,\theta'),\theta') - f(a_1, b_2^*(a_1,\theta'))$$

This leads to the following definition.

Definition 9 (Strong Renegotiation-Proofness)**.** We say that $(f, b_2^*) \in \mathscr{C} \times A_2^{A_1 \times \Theta}$ is *strongly renegotiation-proof* if for all $a_1 \in A_1$ and $\theta \in \Theta$ for which there exists an incentive compatible (g, b_2) such that

$$u_2(a_1, b_2(a_1,\theta),\theta) - g(a_1, b_2(a_1,\theta)) > u_2(a_1, b_2^*(a_1,\theta),\theta) - f(b_2^*(a_1,\theta)) \tag{12}$$

there exists a $\theta' \in \Theta$ such that

$$f(a_1, b_2^*(a_1,\theta')) \geq g(a_1, b_2(a_1,\theta')) \tag{13}$$

and

$$u_2(a_1, b_2(a_1,\theta'),\theta') - g(a_1, b_2(a_1,\theta')) \geq u_2(a_1, b_2^*(a_1,\theta'),\theta') - f(a_1, b_2^*(a_1,\theta')) \tag{14}$$

This is exactly the same as renegotiation-proofness except that it adds condition (14), which allows us to construct beliefs that satisfy intuitive criterion after any renegotiation offer. It can be shown that when we work with this definition, Theorem 1 needs to be modified as follows.

Theorem 4. *(f, b_2^*) is strongly renegotiation-proof if and only if for any $a_1 \in A_1$, $i \in \{1,2,\ldots,n\}$, and $b_2 \in \mathfrak{B}(a_1, i, b_2^*)$ there exists a $k \in \{1,2,\ldots,i-1\}$ such that*

$$\begin{aligned} u_2(a_1, b_2(a_1,\theta^i),\theta^i) - u_2(a_1, b_2^*(a_1,\theta^i),\theta^i) - \min\{0, u_2(a_1, b_2(a_1,\theta^k),\theta^k) - u_2(a_1, b_2^*(a_1,\theta^k),\theta^k)\} \\ + \sum_{j=k}^{i-1} \vec{U}_2(a_1, b_2)_{2j-1} \leq f(a_1)_k - f(a_1)_i \end{aligned} \tag{15}$$

or there exists an $l \in \{i+1, i+2, \ldots, n\}$ such that

$$\begin{aligned} u_2(a_1, b_2(a_1,\theta^i),\theta^i) - u_2(a_1 b_2^*(a_1,\theta^i),\theta^i) - \min\{0, u_2(a_1, b_2(a_1,\theta^l),\theta^l) - u_2(a_1, b_2^*(a_1,\theta^l),\theta^l)\} \\ + \sum_{j=i+1}^{l} \vec{U}_2(a_1, b_2)_{2(j-1)} \leq f(a_1)_l - f(a_1)_i \end{aligned} \tag{16}$$

Proof. Omitted. □

It is also easy to show Propositions 2 and 3 go through with a minor modification similar to the one made in Theorem 4.

5 Equilibrium Outcomes of Games with Contracts

There may be legal or technological constraints that might render contracts non-renegotiable and therefore outcomes that can be supported by non-renegotiable contracts are of interest on their own. Furthermore, understanding non-renegotiable contracts will help place our results within the literature and allow us to isolate the effects of renegotiation. Similarly, and irrespective of whether a contract is renegotiable, there may be valid reasons why a contract maybe observable or unobservable. Legal contracts between a firm and a bank, or a government and an international body, and many compensation contracts are observable yet subject to renegotiation if the parties find it in their benefit to do so. Other contracts can be either secret or subject to renegotiation before the game begins, i.e., they can be unobservable. In this section we will present results regarding the outcomes that can be supported under different assumptions about the contracts.

5.1 Observable Contracts

Let us assume that the contract signed between player 2 and the third-party before the game begins is observable to player 1 but may or may not be renegotiated after player 1 moves in the game.

5.1.1 Non-renegotiable Contracts

If the contracts are observable but not renegotiable, then we can show that player 2 can obtain the best payoff possible given that player 2 plays an increasing strategy and player 1 best responds. More precisely, define the best Stackelberg payoff of player 2 as $\bar{U}_2^B = \max_{b_2 \in B_2^+} \max_{b_1 \in BR_1(b_2)} U_2(b_1, b_2)$ and the worst Stackelberg payoff as $\bar{U}_2^W = \max_{b_2 \in B_2^+} \min_{b_1 \in BR_1(b_2)} U_2(b_1, b_2)$.

Proposition 5. *If contracts are observable, then $\bar{U}_2^B - \delta$ can be supported with non-renegotiable contracts.*

The proof of this result is quite easy. In the definition of the best Stackelberg payoff, player 2 is playing the best increasing strategy, say b_2^*, given that player 1 is playing a best response that is most favorable for player 2. Proposition 1 implies that b_2^* is incentive compatible, i.e., there is a contract, say f^*, that makes it optimal to play. It is easy to show that there is a perfect Bayesian equilibrium of the game with observable and non-renegotiable contracts in which player 2 offers f^* with expected value δ, player 1 plays the most favorable best response to that, say b_1^*, and player 2 plays $b_2^*(b_1^*, \theta)$ after (f^*, b_1^*, θ). Expected payoff of player 2 in such an equilibrium is $\bar{U}_2^B - \delta$.

We can also show that player 2 cannot get a payoff that is smaller than his worst Stackelberg payoff.

Proposition 6. *If contracts are observable, then $\bar{U}_2^W - \delta$ is the smallest payoff that can be supported with non-renegotiable contracts.*

In order to see why let $\hat{b}_{2,a_1}$ $\text{argmin}_{b_2 \in B_2^+} U(a_1, b_2)$ for any $a_1 \in A_1$. In other words, for any a_1, $\hat{b}_{2,a_1}$ is the worst increasing strategy for player 1 that player 2 can play. Since $\hat{b}_{2,a_1}$ is increasing, it can be shown that there is a contract that makes it uniquely optimal to play. Now let $b_1^*(b_2) \in$

$\mathrm{argmin}_{b_1 \in BR_1(b_2)} U_2(b_1, b_2)$, $b_2^* \in \mathrm{argmax}_{b_2 \in B_2^+} U_2(b_1^*(b_2), b_2)$, and $a_1^* = b_1^*(b_2^*)$. Note that $U_2(a_1^*, b_1^*) = \bar{U}_2^W$ and suppose, for contradiction, that player 2 gets a payoff that is strictly smaller than $\bar{U}_2^W - \delta$. We show that there exists a contract that makes it uniquely optimal to play $b_2^*(a_1^*, \theta)$ after a_1^* and $\hat{b}_{2,a_1}(a_1, \theta)$ after any other a_1. If Player 2 offers this contract, player 1 must play a best response to b_2^*. This is because for any $a_1 \notin BR_1(b_1^*)$, we have $U_1(br_1(b_2^*), b_2^*) > U_1(a_1, b_2^*) \geq U_1(a_1, \hat{b}_{2,a_1})$. Therefore, deviation to such a contract yields a gross payoff of at least $U_2(a_1^*, b_1^*) = \bar{U}_2^W$ and a net payoff arbitrarily close to $\bar{U}_2^W - \delta$, a contradiction.

Of course, we have full characterization if player 1's best response correspondence is single-valued, i.e., $BR_1(b_2)$ is a singleton for any $b_2 \in B_2^+$: The unique equilibrium payoff of player 2 that can be supported with observable and non-renegotiable contracts is $\bar{U}_2^B - \delta$.

5.1.2 Renegotiable Contracts

If the contracts are observable and renegotiable, then player 2 can again achieve his Stackelberg payoff, except that the definition of this payoff must reflect the fact that player 2 plays a renegotiation-proof strategy. Define the best and worst renegotiation-proof Stackelberg payoffs of player 2 as $\bar{U}_2^{BR} = \max_{b_2 \in B_2^R} \max_{b_1 \in BR_1(b_2)} U_2(b_1, b_2)$ and the worst Stackelberg payoff as $\bar{U}_2^{WR} = \max_{b_2 \in B_2^R} \min_{b_1 \in BR_1(b_2)} U_2(b_1, b_2)$ and note that the difference in the definitions comes from the fact that player 2 has to play a renegotiation-proof strategy.

Proposition 7. *If contracts are observable, then $\bar{U}_2^{BR} - \delta$ can be supported with renegotiation-proof contracts.*

The proof of this result also constructs an equilibrium in which player 2 receives the best Stackelberg payoff that he can get by playing a renegotiation-proof strategy. There is however a significant complication in the proof compared with the proof of Proposition 5. When contracts are non-renegotiable any deviation from the contract that induces the best Stackelberg outcome under increasing strategies must still induce an increasing strategy. This implies that no deviation can yield a higher payoff. When contracts are renegotiable a deviation may or may not induce a renegotiation-proof strategy and hence we cannot tell whether such a deviation can yield a payoff that is strictly higher than the best Stackelberg payoff that can be obtained by a renegotiation-proof strategy.

What gives us the result is a version of the "renegotiation-proofness principle" that is valid in a class of equilibria.

Definition 10. We say that a perfect Bayesian equilibrium (β, μ) of the game with renegotiable contracts has conservative beliefs if

$$\beta_2(f, a_1, \theta) = g \in \mathscr{C}, \beta_2(f, a_1, \theta, g, y) = b_g(a_1, \theta), \beta_2(f, a_1, \theta, g, n) = b_f(a_1, \theta), \beta_3(I_3(f, a_1, \theta)) = y$$

imply $g(b_g(a_1, \theta)) \geq f(b_f(a_1, \theta))$.

In other words, whenever, in equilibrium, type θ renegotiates the contract from f to g, the third-party should not expect a decrease in the transfer from that type.

Lemma 1 (Renegotiation-Proofness Principle)**.** *Take any a perfect Bayesian equilibrium of the game with renegotiable contracts and assume that it has conservative beliefs. Suppose that a contract f is* *renegotiated after some a_1 and θ. Then, there exists a contract strategy pair that is incentive compatible, renegotiation-proof, and induces the same outcome as f after a_1.*

This lemma tell us that in any equilibrium with conservative beliefs, one can achieve any outcome that is achieved via renegotiation after a_1 by using a renegotiation-proof contract.[17] In the proof of 7 we construct an equilibrium with conservative beliefs that yields $\bar{U}_2^{BR} - \delta$. Lemma 1 implies that there is no deviation that can be obtained via renegotiation, which cannot also be obtained via a renegotiation-proof strategy, and this gives us the desired result.

We can again show that player 2 cannot get a payoff that is smaller than his worst Stackelberg payoff that can be obtained with renegotiation-proof strategies.

Proposition 8. *If contracts are observable, then $\bar{U}_2^{WR} - \delta$ is the smallest payoff that can be supported with non-renegotiable contracts.*

Proof. Proof of this result is similar to that of Proposition 6 and omitted. □

The above results provide sharp predictions for equilibrium outcomes of the games with observable contracts. In particular, they show that third-party contracts play the role of a commitment device to the extent that player 2's strategy respect the constraints brought about by incentive compatibility, in the case of non-renegotiable contracts, and renegotiation-proofness, in the case of renegotiable contracts. The implications of these results in terms of the equilibrium outcomes depend on the specifics of the original game. We present some of these implications our running example and further in Section 6.

EXAMPLE: ENTRY DETERRENCE

Assume that $p \neq 2/3$ so that player 1's best response correspondence is single-valued:

$$br_1(FF) = O, br_1(AA) = E, br_1(FA) = \begin{cases} O, & p > 2/3 \\ E, & p < 2/3 \end{cases}$$

Remember that the set of incentive compatible strategies is $B_2^+ = \{FF, FA, AA\}$ and the set of renegotiation-proof strategies is $B_2^R = \{FA, AA\}$. Therefore, the Stackelberg payoff of player 2 given that he plays an incentive compatible strategy is m, which he achieves by playing FF. Proposition 5 and 6 imply that this is the unique payoff that can be supported with observable and non-renegotiable contracts. In other words, entry-deterrence is the unique equilibrium outcome. How about with RP contracts? If $p > 2/3$, then the Stackelberg payoff is m, obtained by playing FA, whereas if $p < 2/3$, FA does not deter entry and the best player 2 can do in this case is to play AA, with payoff $px + (1-p)z$. In other words, if $p > 2/3$ unique equilibrium outcome is entry-deterrence and if $p < 2/3$ unique equilibrium outcome is entry and accommodate.

5.2 Unobservable Contracts

We now assume that the initial contract between player 2 and the third party is not observable to player 1. Again there are two possibilities: the contract could be renegotiable or non-renegotiable.

[17] This also implies that if there is an outcome $(b_1^*, (b_2^*(\theta))_\theta)$ achieved in an equilibrium with conservative beliefs and via an equilibrium contract that is renegotiated, then there exists an equilibrium in which the same outcome is achieved via a renegotiation-proof contract. In other words, what is called the "renegotiation-proofness" principle in the literature holds as long as we limit ourselves to a certain class of equilibria. Whether it holds for all equilibria is an open question.

5.2.1 Non-Renegotiable Contracts

If contracts are non-renegotiable, we have the following characterization.

Proposition 9. *A strategy profile (b_1^*, b_2^*) of the original game G can be supported with unobservable and non-renegotiable contracts if and only if (b_1^*, b_2^*) is a Bayesian Nash equilibrium of G and b_2^* is increasing in $(\succsim_\theta, \succsim_2)$.*

This result shows that unobservable third-party contracts potentially enlarges the set of outcomes that can arise in equilibrium. Furthermore, while earlier papers showed that, when there is no asymmetric information, any Nash equilibrium of the original game can be supported with unobservable contracts, this result shows that only the subset of Bayesian Nash equilibria in which the second player plays an increasing strategy can be supported if, instead, there is asymmetric information.

This result also has an immediate corollary in terms of the outcomes that can be supported. For any strategy profile $(b_1, b_2) \in A_1 \times A_2^{A_1 \times \Theta}$, we define an *outcome* $(a_1, a_2) \in A_1 \times A_2^\Theta$ of G as $a_1 = b_1$ and $a_2(\theta) = b_2(b_1, \theta)$. Define the individually rational payoff of player 1 as

$$\underline{U}_1^+ = \max_{a_1 \in A_1} \min_{b_2 \in B_2^+} U_1(a_1, b_2). \tag{17}$$

This is the best payoff player 1 can guarantee for herself in game G, given that player 2 plays an increasing strategy.[18] The following easily follows from Proposition 9.

Corollary 1. *An outcome (a_1^*, a_2^*) of the original game G can be supported with unobservable and non-renegotiable contracts if and only if (1) $a_2^*(\theta) \in BR_2(a_1^*, \theta)$ for all θ and (2) $U_1(a_1^*, a_2^*) \geq \underline{U}_1^+$.*

Again, note that, in general, outcomes that are not perfect Bayesian equilibrium outcomes of the original game can also be supported. This can be achieved by writing a contract that leads player 2 to punish player 1 when he deviates from his equilibrium action. Since contracts cannot be conditioned on θ and u_2 has increasing differences, player 2 can only use punishment strategies that are increasing in θ. The best that player 1 can do by deviating is therefore given by $\underline{U}_1$, and his equilibrium payoff cannot be smaller than this payoff. This is condition (2). Condition (1), on the other hand, simply follows from the requirement that only Bayesian Nash equilibrium outcomes can be supported, and hence, player 2 must be best responding along the equilibrium path.

Note that if θ were contractible as well, we would not need to limit the punishment strategies to be increasing. In this case, condition (2) would have the individually rational payoff defined as $\max_{a_1 \in A_1} \min_{b_2 \in A_2^{A_1 \times \Theta}} U_1(a_1, b_2)$. In that case, the result would be the exact analog of those in models without asymmetric information, i.e., Koçkesen and Ok (2004) and Koçkesen (2007).

We should also note that there are interesting environments in which non-contractibility of θ does not restrict the set of outcomes that can be supported with non-renegotiable contracts. For example if player 1's payoff does not depend on θ, then the punishment does not have to depend on θ either. Therefore, one can simply use a constant punishment after each deviation, which would be increasing by construction. A second environment is *games with externalities*, in which u_1 is increasing (or decreasing) in a_2. In this case, after any a_1, the harshest punishment is the lowest (or highest) a_2, which is constant and hence increasing.

[18] We should also note that this is different from the definition of individually rational payoff used in the repeated games literature, which is the minmax payoff rather than the maxmin payoff. The maxmin payoff is at most equal to the minmax payoff.

5.2.2 Renegotiable Contracts

Suppose now that the contracts are unobservable and renegotiable. The counterpart to Proposition 9 is the following:

Proposition 10. *A strategy profile (b_1^*, b_2^*) of the original game G can be supported with unobservable and renegotiation-proof contracts if and only if (b_1^*, b_2^*) is a Bayesian Nash equilibrium of G and b_2^* is increasing and renegotiation-proof.*

This result too has an immediate corollary. Define the best payoff player 1 can guarantee for herself in game G, given that player 2 plays a renegotiation-proof strategy as $\underline{U}_1^R = \max_{a_1 \in A_1} \min_{b_2 \in B_2^R} U_1(a_1, b_2)$. We have the following corollary.

Corollary 2. *An outcome (a_1^*, a_2^*) of the original game G can be supported with unobservable and renegotiation-proof contracts if and only if (1) $a_2^*(\theta) \in BR_2(a_1^*, \theta)$ for all θ and (2) $U_1(a_1^*, a_2^*) \geq \underline{U}_1^R$.*

In order to apply the results on unobservable contracts the crucial piece of information is the individually rational payoff of player 1 given that player 2 plays an increasing or a renegotiation-proof strategy. We illustrate how this can be done in Section 6.1 for a large class of games that we call games with externalities. We show that in those games the only thing that distinguishes the case of renegotiation-proof contracts from non-renegotiable contracts is that the highest type of player 2 must play a best response to any a_1 under renegotiation-proof contracts, whereas the only restriction is that he plays an increasing strategy in the case of non-renegotiable contracts.

EXAMPLE: ENTRY DETERRENCE

Individually rational payoffs of player 1 are given by

$$U_1^+ = \max_{a_1 \in A_1} \min_{b_2 \in B_2^+} U_1(a_1, b_2) = U_1(O, FF) = 0$$

$$U_1^R = \max_{a_1 \in A_1} \min_{b_2 \in B_2^R} U_1(a_1, b_2) = \begin{cases} U_1(O, FA) = 0, & p > 2/3 \\ U_1(E, AA) = 2 - 3p, & p < 2/3 \end{cases}$$

Corollary 1 implies that (O, FF) and (E, AA) can both be supported with unobservable and non-renegotiable contracts. Corollary 2 implies that if $p > 2/3$ both (O, FA) and (E, AA) can be supported with unobservable and RP contracts, whereas if $p < 2/3$ only (E, AA) can be supported.

6 Applications

6.1 Games with Externalities

We say that an original game G is a game with externalities if player 1's payoff is monotonically increasing or decreasing in player 2's action, i.e., for any a_1 and θ, $a_2' \succsim_2 a_2$ implies either $u_1(a_1, a_2', \theta) \geq u_1(a_1, a_2, \theta)$ or $u_1(a_1, a_2', \theta) \leq u_1(a_1, a_2, \theta)$.[19] Such positive or negative externalities are very common in economic models. Indeed, the class of games that satisfy these conditions includes Stackelberg and entry games, sequential Bertrand games with differentiated products, monopolistic screening, and ultimatum bargaining, among others.

[19] Note that player 1's payoff may be increasing in a_2 for some a_1 and decreasing for others.

Assume that contracts are unobservable. Fix $a_1 \in A_1$, let $\underline{a}_2$ ($\overline{a}_2$) be the smallest (largest) element of A_2, and define

$$b_2^+(a_1,\theta) = \begin{cases} \underline{a}_2, & \forall\theta \text{ if } u_1(a_1,a_2,\theta) \text{ increasing in } a_2 \\ \overline{a}_2, & \forall\theta \text{ if } u_1(a_1,a_2,\theta) \text{ decreasing in } a_2 \end{cases}$$

Note that this strategy is increasing in θ and it is the harshest punishment player 2 can inflict upon player 1, i.e., $b_2^+ \in \operatorname{argmin}_{b_2 \in B_2^+} U_1(a_1,b_2)$ for all a_1. In other words, the individually rational payoff of player 1 given that player 2 plays an increasing strategy is given by $\underline{U}_1^+ = \max_{a_1} U_1(a_1,b_2^+)$. We can directly apply Corollary 1 and conclude that an outcome (a_1^*,a_2^*) of the original game G with externalities can be supported with non-renegotiable contracts if and only if $a_2^*(\theta) \in BR_2(a_1^*,\theta)$ for all θ and $U_1(a_1^*,a_2^*) \geq \underline{U}_1^+$.

Which outcomes can be supported with unobservable and renegotiation-proof contracts? Proposition 2 implies that the harshest punishment strategy b_2^+ is not renegotiation proof.[20] Using Proposition 3 we can show that, if u_1 is increasing in a_2, the harshest renegotiation-proof punishment is to make the highest type of player 2 play a best response, while the other types play the smallest a_2 (see Lemma 2 in Section 8). Similarly, if u_1 is decreasing in a_2, the harshest renegotiation-proof punishment is to make the smallest type best respond and the other types play the largest a_2.

More precisely, let $b_2^n(a_1) \in \operatorname{argmin}_{a_2 \in BR_2(a_1,\theta^n)} u_1(a_1,a_2,\theta^n)$ and $b_2^1(a_1) \in \operatorname{argmin}_{a_2 \in BR_2(a_1,\theta^1)} u_1(a_1,a_2,\theta^1)$. Define the punishment strategy as

$$b_2^R(a_1,\theta) = \begin{cases} \underline{a}_2, & \text{if } u_1(a_1,a_2,\theta) \text{ increasing in } a_2 \text{ and } \theta \prec_\theta \theta^n \\ b_2^n(a_1), & \text{if } u_1(a_1,a_2,\theta) \text{ increasing in } a_2 \text{ and } \theta = \theta^n \\ \overline{a}_2, & \text{if } u_1(a_1,a_2,\theta) \text{ decreasing in } a_2 \text{ and } \theta \succ \theta^1 \\ b_2^1(a_1), & \text{if } u_1(a_1,a_2,\theta) \text{ decreasing in } a_2 \text{ and } \theta = \theta^1 \end{cases} \tag{18}$$

The best payoff that player 1 can achieve against this strategy is $\underline{U}_1^R = \max_{a_1} U_1(a_1,b_2^R)$. We can then apply Corollary 2 to games with externalities.

Corollary 3. *An outcome (a_1^*,a_2^*) of an original game with externalities can be supported with unobservable and renegotiation-proof contracts if and only if (1) $a_2^*(\theta) \in BR_2(a_1^*,\theta)$ for all θ and (2) $U_1(a_1^*,a_2^*) \geq \underline{U}_1^R$.*

Therefore, the effect of renegotiation in this environment is very clear. If the contracts are unobservable and non-renegotiable, then any outcome (a_1,a_2) in which player 2 best responds on the equilibrium path and punishes player 1 in the harshest possible way off-the-equilibrium can be supported. With unobservable and renegotiation-proof contracts player 2 cannot punish player 1 in the harshest possible way anymore: the highest (or the lowest) type must play a best response even off-the-equilibrium path.

We next apply these results to a game that has been a canonical example for issues related to commitment: Stackelberg and entry-deterrence games. This example will also give us an opportunity to discuss the implications of observable contracts.

[20] See Lemma 3 in Section 8, which shows that renegotiation-proofness of a strategy $b_2 \in A_2^{A_1 \times \Theta}$ implies that the highest (lowest, resp.) type does not have a profitable deviation to a higher (lower, resp.) action.

EXAMPLE: QUANTITY COMPETITION AND ENTRY-DETERRENCE

Consider a Stackelberg game in which firm 1 moves first by choosing an output level $q_1 \in Q_1$ and firm 2, after observing q_1, chooses its own output level $q_2 \in Q_2$. Inverse demand function is given by $P(q_1, q_2) = \max\{0, \alpha - q_1 - q_2\}$, where $\alpha > 0$, and we assume Q_i a rich enough finite subset of $\mathbb{R}_+$ whose largest element is α.[21] Cost function of firm 1 is $C_1(q_1) = cq_1$, where c is common knowledge, whereas the cost function of firm 2 is $C_2(q_2) = \theta q_2$. We assume that $\theta \in \{\theta^1, \theta^2, \ldots, \theta^n\}$, where $n \geq 2$, is private information of firm 2 and $\theta^1 < \theta^2 < \cdots < \theta^n$. Firm 1 believes that the probability of θ^i is given by $p(\theta^i)$ and for ease of exposition we assume that expected value of θ is equal to c. The profit function of firm i is given by $\pi_i(q_1, q_2, \theta) = P(q_1, q_2)q_i - C_i(q_i)$ and we assume that both firms are profit maximizers.

To ensure positive output levels in equilibrium we assume that $\alpha > 2\theta^n - c$, in which case the (Stackelberg) equilibrium outcome of this game is given by

$$\left(q_1^s, q_2^s(\theta)\right) = \left(\frac{\alpha - c}{2}, \frac{\alpha - 2\theta + c}{4}\right)$$

Define the game G as follows: Let $A_1 = Q_1$ and $A_2 = \{-q_2 : q_2 \in Q_2\}$ and define $\succsim_i$ on A_i as $a_i \succsim_i a_i' \Leftrightarrow a_i \geq a_i'$ and $\succsim_\theta$ as $\theta \succsim_\theta \theta' \Leftrightarrow \theta \geq \theta'$. Let the payoff function of player i be given by $u_i(a_1, a_2, \theta) = \pi_i(a_1, -a_2, \theta)$, for any $(a_1, a_2) \in A_1 \times A_2$. The game G is strategically equivalent to the Stackelberg game defined in the previous paragraph. It is also easy to show that u_2 has strictly increasing differences in (a_2, θ) and u_1 is increasing in a_2.

Let us first assume that contracts are unobservable. Since u_2 has strictly increasing differences and u_1 is increasing in a_2, we can apply Corollary 1 and Corollary 2 to characterize all the outcomes that can be supported with non-renegotiable as well as renegotiation-proof third-party contracts. In order to apply Corollary 1, we need to calculate the individually rational payoff of player 1, i.e., $\underline{U}_1^+$ as defined in equation (17). The harshest punishment firm 2 can inflict is to drive the price down to zero by producing α for any type θ. Since this is a constant (and hence an increasing) strategy, it follows that $\underline{U}_1 = 0$. In other words, any outcome $(a_1^*, a_2^*(\theta))$ such that firm 2 best responds to a_1^* and firm 1 gets at least zero profit can be supported with non-renegotiable contracts. In particular, entry can be deterred with non-renegotiable contracts.

Can entry be deterred with renegotiation-proof contracts? In order to apply Corollary 2, we need to first calculate player 1's individually rational payoff given that player 2 plays a renegotiation-proof strategy. The discussion above implies that the harshest punishment is obtained when the highest type of player 2 best responds while the other types choose the lowest a_2, i.e., $a_2 = -\alpha$. Player 1's expected payoff when player 2 plays this strategy is given by $\frac{1}{2}p(\theta^n)\,(\alpha + \theta^n - a_1)\,a_1 - ca_1$. Its maximum, i.e., player 1's individually rational payoff, is therefore equal to

$$\underline{U}_1^R = \begin{cases} 0, & p(\theta^n)(\alpha + \theta^n) - 2c \leq 0 \\ \frac{(p(\theta^n)(\alpha+\theta^n)-2c)^2}{8p(\theta^n)}, & p(\theta^n)(\alpha + \theta^n) - 2c > 0 \end{cases}$$

Condition (1) of Corollary 2 requires that $a_2^*(\theta) = \frac{a_1^* + \theta - \alpha}{2}$ for all θ, and hence $U_1(a_1^*, a_2^*) = \frac{1}{2}(\alpha - c - a_1^*)a_1^*$. Therefore, by condition (2), any outcome such that $\frac{1}{2}(\alpha - c - a_1^*)a_1^* \geq \underline{U}_1^R$ can be supported.

Also note that if $p(\theta^n)(\alpha + \theta^n) - 2c > 0$, then $\underline{U}_1^R$ is strictly positive. This implies that entry cannot be deterred if $p(\theta^n)(\alpha + \theta^n) - 2c > 0$. Therefore, we have the following result:

[21] We introduce this assumption so that player 2 can choose a high enough output level to drive the price to zero.

Corollary 4. *Entry can be deterred with unobservable and non-renegotiable contracts. It can be deterred with unobservable and renegotiation-proof contracts if and only if $p(\theta^n)(\alpha+\theta^n)-2c \le 0$.*

Now let us assume that contracts are observable. The best payoff that player 2 can obtain in the original game is the monopoly outcome, i.e., $a_1^* = 0$ and $a_2^*(\theta) = \frac{a_1^*+\theta-\alpha}{2}$. If contracts are non-renegotiable, then Player 2 can obtain this outcome exactly the same way as with unobservable contracts: If player 1 plays any $a_1 > 0$, punish him by flooding the market, i.e., choose $a_2 = -\alpha$. In other words, with observable and non-renegotiable contracts the unique outcome is the monopoly (entry-deterrence) outcome.

Could player 2 achieve the monopoly outcome with renegotiation-proof contracts? The above analysis implies that the answer is yes as long as $\underline{U}_1^R = 0$, i.e., $p(\theta^n)(\alpha+\theta^n)-2c \le 0$. It is easy to see that if this condition holds, then the unique equilibrium outcome that can be achieved with observable and renegotiation-proof contracts is the monopoly outcome. If, on the other hand, $p(\theta^n)(\alpha+\theta^n)-2c > 0$, then monopoly outcome can no longer be supported with renegotiation-proof contracts. However, player 2 can obtain the following outcome: Player 1 plays a_1^*, where a_1^* is the smallest a_1 such that

$$\frac{1}{2}(\alpha - c - a_1^*)a_1^* \ge \frac{\big(p(\theta^n)(\alpha+\theta^n)-2c\big)^2}{8p(\theta^n)}$$

and player 2 plays $a_2^*(\theta) = \frac{a_1^*+\theta-\alpha}{2}$ for all θ. In this outcome, player 1 produces the smallest amount consistent with player 2 punishing with the harshest possible renegotiation-proof strategy off-the-equilibrium and best responding on the equilibrium path.

Dewatripont (1988) has also analyzed a similar entry game and showed that entry can be deterred with renegotiation-proof contracts under certain conditions. His conditions are different from ours because he uses a different renegotiation-proofness concept, namely durability, first introduced by Holmstrom and Myerson (1983). A decision rule is durable if and only if the parties involved would never unanimously approve a change from this decision rule to any other decision rule. Holmstrom and Myerson also show that this is equivalent to interim incentive efficiency when there is only one player with private information. In our context, only player 2 has private information and hence a contract-strategy pair (f, b_2^*) is interim incentive efficient (and therefore durable) if and only if there is no $a_1 \in A_1$ and an incentive compatible (g, b_2) such that after a_1 every type of player 2 and the third-party do better under (g, b_2), with at least one doing strictly better.

We have a characterization of durable strategies for the two-type case, i.e., when $\Theta = \{\theta^1, \theta^2\}$, and even in that case, the relationship between our concept of renegotiation-proofness and durability turns out to be quite subtle. It is not difficult to show that neither concept implies the other one in general. However, in games with externalities it can be shown that durability implies renegotiation-proofness. The entry-deterrence game is a game with externalities, and therefore, if entry can be deterred with durable contracts, it can also be deterred with renegotiation-proof contracts. In fact, in the entry-deterrence game player 2's payoff function is single-peaked and for such environments we have a complete characterization of durable outcomes that is particularly easy to apply. Using this characterization, we can show that the relationship between durability and renegotiation-proofness is strict.

Proposition 11. *In the entry-deterrence game with two types, if $p_1(\theta^2+\alpha) > (\theta^2-\theta^1)$, then entry can* *be deterred with renegotiation-proof contracts but not with durable contracts.*

Proof. Omitted. ☐

Remember that the harshest renegotiation-proof punishment strategy of the incumbent is to flood the market if entry occurs, except for the highest type (type θ^2), who has to best respond. Durability still requires that the highest type best responds. The difference is that flooding the market for type θ^1 is not a durable strategy: There is a restriction on how much the incumbent can produce in response to entry, which is condition (d) of Proposition 1 in Dewatripont (1988).

7 Conclusion

In this paper we characterized incentive compatible and renegotiation-proof third-party contracts and strategies in extensive form games with incomplete information. We applied our results to two-stage games and showed that when the contracts are observable to the first mover, then the second mover obtains her Stackelberg payoff that can be achieved with renegotiation-proof strategies. When the contracts are not observable, then some kind of a "folk theorem" is true: Any outcome in which the second mover best responds to the first mover's action and the first mover obtains his individually rational payoff can be supported. In the definition of the individually rational payoff, player 2 is restricted to using increasing and renegotiation-proof strategies. The restriction imposed by renegotiation-proofness is particularly transparent in games with externalities, i.e., games in which the first mover's payoff is monotonically increasing (or decreasing) in the second mover's action. In this class of games player 2 can induce player 1 to play player 2's favorite action by punishing him if he plays some other action. If player 1's payoff is increasing in player 2's action, then the worst punishment is to play the lowest possible action for every type of player 2. However, this is not a renegotiation-proof strategy. The worst renegotiation-proof punishment is to best respond for the highest type while the others play the smallest action.

Overall, we conclude that even with renegotiation-proof contracts, one can support outcomes that are not perfect Bayesian equilibrium outcomes of the original game, and this may benefit the second mover in many games, such as the entry game.

8 Proofs

In the game with non-renegotiable contracts, player 2 has an information set at the beginning of the game, which we identify with the null history $\emptyset$, and an information set for each $(f,\theta,a_1) \in \mathscr{C} \times \Theta \times A_1$, where $\mathscr{C} = \mathbb{R}^{A_1 \times A_2}$. Player 3 has an information set for each $f \in \mathscr{C}$. If contracts are unobservable, then player 1 has only one information set, given by $\mathscr{C}$. If contracts are observable, then player 1 has an information set for each $f \in \mathscr{C}$. In the game with renegotiable contracts, player 2 has additional information sets corresponding to each history (f,θ,a_1,g,y) and (f,θ,a_1,g,n) and player 3 has an additional information set of each (f,a_1,g), which we denote by $I_3(f,a_1,g)$.

We first introduce some notation. Let the number of elements in Θ be equal to n and order its elements so that $\theta^n \succsim_\theta \theta^{n-1} \succsim_\theta \cdots \theta^2 \succsim_\theta \theta^1$. Let e_i be the i^{th} standard basis row vector for $\mathbb{R}^n$ and define the row vector $d_i = e_i - e_{i+1}, i = 1,2,\ldots,n-1$. Let D be the $2(n-1)\times n$ matrix whose row $2i-1$ is d_i and row $2i$ is $-d_i$, $i = 1,\ldots,n-1$. For any $a_1 \in A_1$ and $b_2 \in A_2^{A_1\times\Theta}$ define $\vec{U}_2(a_1,b_2)$ as a column vector with $2(n-1)$ components, where component $2i-1$ is given by $u_2(a_1, b_2(a_1,\theta^i),\theta^i) - u_2(a_1, b_2(a_1,\theta^{i+1}),\theta^i)$ and component $2i$ is given by $u_2(a_1, b_2(a_1,\theta^{i+1}),\theta^{i+1}) - u_2(a_1, b_2(a_1,\theta^i),\theta^{i+1})$, $i = 1,2,\ldots,n-1$.

Notation 1. Given two vectors $x, y \in R^n$

1. $x \geq y$ if and only if $x_i \geq y_i$, for all $i = 1,2,\ldots,n$;
2. $x > y$ if and only if $x_i \geq y_i$, for all $i = 1,2,\ldots,n$ and $x \neq y$;
3. $x \gg y$ if and only if $x_i > y_i$, for all $i = 1,2,\ldots,n$.

Similarly for $\leq$, $<$, and $\ll$.

For any $a_1 \in A_1$, $b_2 \in A_2^{A_1\times\Theta}$ and $f \in \mathscr{C}$, let $f(a_1, b_2)$ be the column vector with n components, where i^{th} component is given by $f(a_1, b_2(a_1,\theta^i)), i = 1,2,\ldots,n$. For any strategy profile (b_1, b_2) of G define the expected transfer from player 2 to the third-party as $F(b_1, b_2) = \sum_{\theta\in\Theta} p(\theta) f(b_1, b_2(b_1,\theta))$.

Proof of Proposition 5. Let $b_2^* \in \operatorname{argmax}_{b_2\in B_2^+} \max_{b_1\in BR_1(b_2)} U_2(b_1, b_2)$ and $b_1^* \in \operatorname{argmax}_{b_1\in BR_1(b_2^*)} U_2(b_1, b_2^*)$. Note that $\bar{U}_2^B = U_2(b_1^*, b_2^*)$. Since b_2^* is increasing by construction, there exists a contract f^* such that (f^*, b_2^*) is incentive compatible and $F(b_1^*, b_2^*) = \delta$. For any $f \in \mathscr{C}, a_1 \in A_1, \theta \in \Theta$ choose $b_{2,f} \in \operatorname{argmax}_{a_2\in A_2} u_2(a_1, a_2,\theta) - f(a_1, a_2)$ and $b_{1,f} \in \operatorname{argmax}_{a_1'\in A_1} U_1(a_1', b_{2,f})$.

Consider the following assessment (β,μ) of $\Gamma(G)$: $\beta_2(\emptyset) = f^*$, $\beta_3(f^*) = y$, $\beta_3(f) = y$ iff $F(b_{1,f}, b_{2,f}) \geq \delta$, $\beta_1(f^*) = b_1^*$, $\beta_1(f) = b_{1,f}$, for $f \neq f^*$, $\beta_2(f^*,\theta, a_1) = b_2^*(a_1,\theta)$, $\beta_2(f,\theta, a_1) = b_{2,f}(a_1,\theta)$ for all $f \neq f^*$, $a_1 \in A_1$, and $\theta \in \Theta$.

If player 2 offers any contract $f \neq f^*$, the continuation play will be $(b_{1,f}, b_{2,f})$. If $F(b_{1,f}, b_{2,f}) < \delta$ it will be rejected and hence it cannot be a profitable deviation. If $F(b_{1,f}, b_{2,f}) \geq \delta$, then

$$U_2(b_1^*, b_2^*) - F(b_1^*, b_2^*) = U_2(b_1^*, b_2^*) - \delta \geq U_2(b_{1,f}, b_{2,f}) - F(b_{1,f}, b_{2,f})$$

by construction. Therefore, it is optimal for player 2 to offer f^*. Sequential rationality at other information sets are easily checked and we conclude that this assessment is a perfect Bayesian equilibrium of the game with observable contracts. □

Proof Proposition 6. Let $b_1^*(b_2) \in \operatorname{argmin}_{b_1\in BR_1(b_2)} U_2(b_1, b_2)$, $b_2^* \in \operatorname{argmax}_{b_2\in B_2^+} U_2(b_1^*(b_2), b_2)$, and $a_1^* = b_1^*(b_2^*)$. Note that $U_2(a_1^*, b_1^*) = \bar{U}_2^W$ and suppose, for contradiction, that player 2 gets a payoff $\tilde{U}_2 < \bar{U}_2^W - \delta$. We will show that player 2 can offer a contract that supports (a_1^*, b_2^*) and yields a higher payoff.

For any a_1 choose $\hat{b}_{2,a_1} \in \operatorname{argmin}_{b_2\in B_2^+} U_1(a_1, b_2)$. By construction $\hat{b}_{2,a_1}$ is increasing and hence there exists a contract that makes it optimal to play. We will further show that there exists a contract that makes it the unique optimal strategy after a_1. Assume without loss of generality that $\hat{b}_{2,a_1}(a_1,\theta) \neq \hat{b}_{2,a_1}(a_1,\theta')$ whenever $\theta \neq \theta'$ and hence $\hat{b}_{2,a_1}(a_1,\theta^i) \succ_\theta \hat{b}_{2,a_1}(a_1,\theta^{i-1})$ for all $i = 1,\ldots,n$.[22] Define $\vec{U}_2(a_1, \hat{b}_{2,a_1})$ as usual and note that strictly increasing differences and $\hat{b}_{2,a_1}(a_1,\theta^i) \succ_\theta \hat{b}_{2,a_1}(a_1,\theta^{i-1})$ imply that

$$\vec{U}_2(a_1, \hat{b}_{2,a_1})_{2i-1} + \vec{U}_2(a_1, \hat{b}_{2,a_1})_{2i} > 0, \quad \forall i = 1,\ldots,n-1.$$

[22] If there exist i such that $\hat{b}_{2,a_1}(a_1,\theta^i) = \hat{b}_{2,a_1}(a_1,\theta^{i-1})$ simply eliminate the incentive compatibility constraint between them and set $\hat{f}_{a_1}(a_1, \hat{b}_{2,a_1}(a_1,\theta^i)) = \hat{f}_{a_1}(a_1, \hat{b}_{2,a_1}(a_1,\theta^{i-1}))$.

We will show that there exists f^{a_1} such that $Df^{a_1} \ll \vec{U}_2(a_1, \hat{b}_{2,a_1})$. Define

$$A = \begin{pmatrix} \vec{U}_2(a_1, \hat{b}_2) & -D \\ 1 & 0 \end{pmatrix}$$

and note that there exists f^{a_1} such that $Df^{a_1} \ll \vec{U}_2(a_1, \hat{b}_{2,a_1})$ iff there exists x such that $Ax \gg 0$.[23] By Gordan's Theorem, this is true iff $A'y = 0$ implies $y \le 0$. It is easy to show that $A'y = 0$ implies $y_1 = y_2, y_3 = y_4, \cdots, y_{2(n-1)-1} = y_{2(n-1)}$. Therefore,

$$A'y = y_{2(n-1)+1} + \sum_{i=1}^{n-1} (\vec{U}_2(a_1, \hat{b}_{2,a_1})_{2i-1} + \vec{U}_2(a_1, \hat{b}_{2,a_1})_{2i}) y_{2i-1}$$

$\vec{U}_2(a_1, \hat{b}_{2,a_1})_{2i-1} + \vec{U}_2(a_1, \hat{b}_{2,a_1})_{2i} > 0, \forall i = 1, \ldots, n-1$, and $A'y = 0$ imply $y \le 0$.

Let $\varepsilon > 0$ be small and define $f(b_1^*, a_2) = \delta + \varepsilon$ for all a_2. For any $a_1 \neq b_1^*$ define

$$f(a_1, a_2) = \begin{cases} f_i^{a_1}, & a_2 = \hat{b}_{2,a_1}(a_1, \theta^i) \\ \infty, & \text{otherwise} \end{cases}$$

Under this contract, player 2 plays a best response to a_1^* and according to $\hat{b}_{2,a_1}$ after any $a_1 \neq a_1^*$. Player 1, on the other hand, must play a best response to b_2^*. This is because for any $a_1 \notin BR_1(b_1^*)$, we have $U_1(br_1(b_2^*), b_2^*) > U_1(a_1, b_2^*) \ge U_1(a_1, \hat{b}_{2,a_1})$. Therefore, deviation to such a contract yields a payoff of $U_2(b_1^*, b_2^*) - \delta - \varepsilon > \tilde{U}_2$, for small enough ε. In other words, player 2 has a profitable deviation, contradicting that $\tilde{U}_2$ is an equilibrium payoff. $\square$

Proof of Lemma 1. Fix a a perfect Bayesian equilibrium with conservative beliefs and suppose that contract f is renegotiated after some a_1 and θ. Let the set of types after which f is renegotiated be Θ^R and $\Theta^{NR} = \Theta \setminus \Theta^R$. For any $\theta \in \Theta^R$, let g_θ be the new contract and $b_{g_\theta}(a_1, \theta)$ be the new strategy of player 2 after a_1 and θ. Also let $b_f(a_1, \theta)$ be the equilibrium strategy of player 2 after a_1 and θ if he does not renegotiate f. In other words, we have $\beta_2(f, \theta, a_1) = b_f(a_1, \theta)$, $\forall \theta \in \Theta^{NR}, \beta_2(f, \theta, a_1) = g_\theta$, $\forall \theta \in \Theta^R$, $\beta_2(f, \theta, a_1, g_\theta, y) = b_{g_\theta}(a_1, \theta)$, $\beta_2(f, \theta, a_1, g_\theta, n) = b_f(a_1, \theta)$, and $\beta_3(I_3(f, a_1, g_\theta)) = y$. For ease of exposition we will omit the reference to a_1 in the following. Consider the following mixture menu:

$$\{(g_\theta(b_{g_\theta}(\theta)), b_{g_\theta}(\theta))_{\theta \in \Theta^R}\} \cup \{(f(b_f(\theta)), b_f(\theta))_{\theta \in \Theta^{NR}}\}$$

It is clear that this menu replicates the outcome induced by f after a_1. We also claim that this menu is incentive compatible and renegotiation proof after a_1.

Incentive compatibility of (f, b_f) implies that no two types in Θ^{NR} has an incentive to mimic each other. Consider $\theta, \theta' \in \Theta^R$ and suppose, for contradiction, that

$$u_2(b_{g_\theta}(\theta), \theta) - g_\theta(b_{g_\theta}(\theta)) < u_2(b_{g_{\theta'}}(\theta'), \theta) - g_{\theta'}(b_{g_{\theta'}}(\theta'))$$

But then type θ could increase her payoff after (f, a_1) by offering $g_{\theta'}$ and playing $b_{g_{\theta'}}(\theta')$ rather than

[23] To see this let $x = \begin{pmatrix} \zeta \\ f^{a_1} \end{pmatrix}$

offering g_θ and playing $b_{g_\theta}(\theta)$.[24]

Now let $\theta' \in \Theta^{NR}$ and $\theta \in \Theta^R$, and suppose for contradiction that

$$u_2(b_f(\theta'),\theta') - f(b_f(\theta')) < u_2(b_{g_\theta}(\theta),\theta') - g_\theta(b_{g_\theta}(\theta))$$

This implies that after (f, a_1) offering g_θ, which is accepted in equilibrium, and playing $b_{g_\theta}(\theta)$ is a profitable deviation for type θ'.

Finally, let $\theta' \in \Theta^{NR}$ and $\theta \in \Theta^R$ and suppose, for contradiction, that

$$u_2(b_{g_\theta}(\theta),\theta) - g_\theta(b_{g_\theta}(\theta)) < u_2(b_f(\theta'),\theta) - f(b_f(\theta'))$$

But then type θ could play $b_f(\theta')$ after (f, a_1) and receive a higher payoff rather than offering g_θ, which is accepted, and playing $b_{g_\theta}(\theta)$. This proves that the mixture menu is incentive compatible.

Suppose now, for contradiction, that the mixture menu is not renegotiation-proof after a_1. Then, there exists θ and an incentive compatible contract strategy pair (h, b_h) such that if $\theta \in \Theta^{NR}$, then,

$$u_2(b_h(\theta),\theta) - h(b_h(\theta)) > u_2(b_f(\theta),\theta) - f(b_f(\theta)) \tag{19}$$

if $\theta \in \Theta^R$, then

$$u_2(b_h(\theta),\theta) - h(b_h(\theta)) > u_2(b_{g_\theta}(\theta),\theta) - g_\theta(b_{g_\theta}(\theta)) \tag{20}$$

and

$$h(b_h(\hat\theta)) > f(b_f(\hat\theta)), \forall \hat\theta \in \Theta^{NR} \tag{21}$$

$$h(b_h(\hat\theta)) > g_{\hat\theta}(b_{g_{\hat\theta}}(\hat\theta)), \forall \hat\theta \in \Theta^R \tag{22}$$

Since $g_{\hat\theta}$ is accepted for all $\hat\theta \in \Theta^R$ and the equilibrium has conservative beliefs,

$$g_{\hat\theta}(b_{g_{\hat\theta}}(\hat\theta)) \geq f(b_f(\hat\theta)),\ \forall \hat\theta \in \Theta^R \tag{23}$$

which, together with (21) and (22), implies that

$$h(b_h(\hat\theta)) > f(b_f(\hat\theta)), \forall \hat\theta \in \Theta \tag{24}$$

Suppose first that $\theta \in \Theta^{NR}$. Inequalities (19) and (24) imply that after (f, a_1) type θ could offer h, which would be accepted, and increase her payoff, a contradiction that in equilibrium she plays $b_f(\theta)$ after (f, a_1).

Similarly, if $\theta \in \Theta^R$, then (20) and (24) imply that after (f, a_1) type θ could offer h, which would be accepted, and increase her payoff, rather than offering g_θ, a contradiction. Therefore, the mixture menu is renegotiation-proof.

Since the mixture is incentive compatible we can easily extend it to a contract whose domain is

[24]Note that $g_{\theta'}$ is accepted after (f, a_1) in equilibrium by assumption.

the entire A_2 rather than just the range of b_f and b_{g_θ}. Define the new contract as

$$h(a_2) = \begin{cases} f(a_2), & \exists\theta : a_2 = b_f(\theta) \\ g_\theta(a_2), & \exists\theta : a_2 = b_{g_\theta}(\theta) \\ \infty, & \text{otherwise} \end{cases}$$

and note that h is well-defined since incentive compatibility of the mixture menu implies that whenever $b_f(\theta') = b_{g_\theta}(\theta) = a_2$ for some $\theta \in \Theta^R$ and $\theta' \in \Theta^{NR}$ we must also have $f(a_2) = g_\theta(a_2)$. □

Proof of Proposition 7. Let $b_2^* \in \operatorname{argmax}_{b_2 \in B_2^R} \max_{b_1 \in BR_1(b_2)} U_2(b_1, b_2)$ and $b_1^* = \operatorname{argmax}_{b_1 \in BR_1(b_2^*)} U_2(b_1, b_2^*)$. Note that $\bar{U}_2^{BR} = U_2(b_1^*, b_2^*)$. Since b_2^* is increasing and renegotiation-proof, there exists $f^* \in \mathscr{C}$ such that (f^*, b_2^*) is incentive compatible and renegotiation-proof with $F^*(b_1^*, b_2^*) = \delta$. For any $f \in \mathscr{C}$, a_1, and θ, let $b_{2,f}(a_1,\theta) \in \operatorname{argmax}_{a_2} u_2(a_1, a_2, \theta) - f(a_1, a_2)$ and $g_{(f,\theta,a_1)} \in \operatorname{argmax}_g u_2(a_1, b_{2,g}(a_1,\theta),\theta) - g(a_1, b_{2,g}(a_1,\theta))$ subject to $g(a_1, b_{2,g}(a_1,\theta')) \geq f(a_1, b_{2,f}(a_1,\theta'))$ for all θ'.

Consider the following assessment (β,μ): $\beta_2(\emptyset) = f^*$; $\beta_1(f^*) = b_1^*$, $\beta_3(f^*) = y$, $\beta_2(f^*,\theta,a_1) = b_2^*(a_1,\theta)$ for all (a_1,θ);

$$\beta_2(f,\theta,a_1) = \begin{cases} g_{(f,\theta,a_1)}, \text{ if} & \begin{array}{r} u_2(a_1, b_{2,g_{(f,\theta,a_1)}}(a_1,\theta),\theta) - g_{(f,\theta,a_1)}(a_1, b_{2,g_{(f,\theta,a_1)}}(a_1,\theta)) \\ > u_2(a_1, b_{2,f}(a_1,\theta),\theta) - f(a_1, b_{2,f}(a_1,\theta)) \end{array} \\ b_{2,f}(a_1,\theta), & \text{otherwise} \end{cases}$$

for any $f \neq f^*$ and (θ, a_1); $\beta_2(f,\theta,a_1,g,y) = b_{2,g}(a_1,\theta)$ and $\beta_2(f,\theta,a_1,g,n) = b_{2,f}(a_1,\theta)$ for all $f \neq f^*$ and (a_1,θ,g); $\beta_2(f^*,\theta,a_1,g,n) = b_2^*(a_1,\theta)$ for all (a_1,θ,g);

$$\beta_3(I_3(f^*,a_1,g)) = \begin{cases} y, & g(a_1, b_{2,g}(a_1,\theta)) > f^*(a_1, b_2^*(a_1,\theta)) \quad \forall\theta \\ n, & \text{otherwise} \end{cases}$$

and

$$\beta_3(I_3(f,a_1,g)) = \begin{cases} y, & \text{if } g(a_1, b_{2,g}(a_1,\theta)) \geq f(a_1, b_{2,f}(a_1,\theta)) \quad \forall\theta \\ n, & \text{otherwise} \end{cases}$$

for any a_1, g and $f \neq f^*$. Obviously, any $f \neq f^*$ induces a continuation strategy b_2^f for player 2, which may involve renegotiation after some θ. Let player 1 play the same best response to the continuation play irrespective of the contract that induces it. Let the third-party accept f iff continuation play yields expected transfers at least equal to δ. Specify beliefs as follows: $\mu(I_3(f^*,a_1,g))(\theta) = p(\theta)$ if $g(a_1, b_{2,g}(a_1,\theta)) > f^*(a_1, b_2^*(a_1,\theta))$ for all θ and $\mu(I_3(f^*,a_1,g))(\theta') = 1$ if there exists θ' such that $f^*(a_1, b_2^*(a_1,\theta')) \geq g(a_1, b_{2,g}(a_1,\theta'))$; For any $f \neq f^*$ and (a_1,g), $\mu(I_3(f,a_1,g))(\theta) = p(\theta)$ if $g(a_1, b_{2,g}(a_1,\theta)) \geq f(a_1, b_{2,f}(a_1,\theta))$ for all θ and $\mu^*(I_3(f,a_1,g))(\theta') = 1$ if there exists θ' such that $f(a_1, b_{2,f}(a_1,\theta')) > g(a_1, b_{2,g}(a_1,\theta'))$.

Now consider any contract $f \neq f^*$. If $(f, b_{2,f})$ is renegotiation-proof, then $b_{2,f} \in B_2^R$ and hence f cannot yield a higher payoff than f^*. Therefore, suppose that $(f, b_{2,f})$ is not renegotiation-proof and let b_2^f be the induced strategy, which includes renegotiation after some a_1 and θ. Since $\beta_3(I_3(f,a_1,g)) = y$ iff $g(a_1, b_{2,g}(a_1,\theta)) \geq f(a_1, b_{2,f}(a_1,\theta))$ for all $\theta \in \Theta$, the equilibrium constructed above has conservative beliefs. Lemma 1 therefore implies that there exists $(h, b_{2,h})$ which is incentive compatible and renegotiation-proof and induces the same outcome as (f, b_2^f). But no renegotiation-proof strategy can yield a payoff that is higher than $\bar{U}_2^{BR}$ and hence deviation to f cannot be profitable.

Sequential rationality at other information sets and consistency of beliefs can be checked easily to show that the above assessment is a perfect Bayesian equilibrium. □

Proof. (Proof of Proposition 9) **(Only if)** Suppose that (b_1^*, b_2^*) can be supported. Then, there exists a perfect Bayesian equilibrium (β^*, μ^*) that induces (b_1^*, b_2^*), i.e., $\beta_2^*(\emptyset) = f^*$, $\beta_3(f^*) = y$, $\beta_1^*(\mathscr{C}) = b_1^*$, $\beta_2^*(f^*, \theta, a_1) = b_2^*(a_1, \theta)$ for all $a_1 \in A_1$ and $\theta \in \Theta$. The fact that (b_1^*, b_2^*) is a Bayesian Nash equilibrium of G is a direct consequence of sequential rationality of players 1 and 2. It must also be the case that it is optimal to play according to b_2^* under f^*. Increasing differences and Proposition 1 implies that b_2^* is increasing.

[If] Let (b_1^*, b_2^*) be a Bayesian Nash equilibrium of G such that b_2^* is increasing. Proposition 1 implies that there exists a contract f' such that (f', b_2^*) is incentive compatible. It is not difficult to show that we can find such a contract whose expected value under (b_1^*, b_2^*) is equal to δ. So assume $F'(b_1^*, b_2^*) = \delta$. For any $b_2 \in A_2^{A_1 \times \Theta}$ and $a_1 \in A_1$, let $b_2(a_1, \Theta)$ be the image of Θ under $b_2(a_1, .)$. Define

$$f^*(a_1, a_2) = \begin{cases} f'(a_1, a_2), & \text{if } a_2 \in b_2^*(a_1, \Theta) \\ \infty, & \text{otherwise} \end{cases}$$

for any $(a_1, a_2) \in A_1 \times A_2$, and

$$b_{2,f}^*(a_1, \theta) = \begin{cases} b_2^*(a_1, \theta), & f = f^* \\ \in \operatorname{argmax}_{a_2} u_2(a_1, a_2, \theta) - f(a_1, a_2), & f \neq f^* \end{cases}$$

for any $f \in \mathscr{C}$, $a_1 \in A_1$, and $\theta \in \Theta$. Consider the assessment (β^*, μ^*) of $\Gamma(G)$, where $\beta_2^*[\emptyset] = f^*$, $\beta_3[f] = y$ iff $F(b_1^*, b_{2,f}^*) \geq \delta$, $\beta_1^*[\mathscr{C}] = b_1^*$, $\beta_2^*[f, \theta, a_1] = b_{2,f}^*(a_1, \theta)$ for all $f \in \mathscr{C}$, $a_1 \in A_1$, and $\theta \in \Theta$, and $\mu^*[\mathscr{C}](f^*) = 1$. It is easy to check that this assessment induces (b_1^*, b_2^*) and is a perfect Bayesian equilibrium of $\Gamma(G)$. □

Proof of Proposition 10. **[If]** Let (b_1^*, b_2^*) be a Bayesian Nash equilibrium of G such that b_2^* is increasing and renegotiation-proof. This implies that there exists $f' \in \mathscr{C}$ such that (f', b_2^*) is incentive compatible and renegotiation-proof. Let $f^*(a_1, a_2) = f'(a_1, a_2) - F'(b_1^*, b_2^*) + \delta$ for all (a_1, a_2) and note that $F^*(b_1^*, b_2^*) = \delta$. Furthermore, using Theorem 1, it can be easily checked that (f^*, b_2^*) is incentive compatible and renegotiation-proof. For any $f \in \mathscr{C}$, a_1, and θ, let $b_{2,f}(a_1, \theta) \in \operatorname{argmax}_{a_2} u_2(a_1, a_2, \theta) - f(a_1, a_2)$ and $g_{(f,\theta,a_1)} \in \operatorname{argmax}_g u_2(a_1, b_{2,g}(a_1, \theta), \theta) - g(a_1, b_{2,g}(a_1, \theta))$ subject to $g(a_1, b_{2,g}(a_1, \theta')) \geq f(a_1, b_{2,f}(a_1, \theta'))$ for all θ'.

Consider the following assessment (β^*, μ^*) of $\Gamma_R(G)$: $\beta_2^*(\emptyset) = f^*$; $\beta_3(f) = y$ iff continuation play yields an expected transfer of at least δ, $\beta_1^*(\mathscr{C}) = b_1^*$, $\beta_2^*(f^*, \theta, a_1) = b_2^*(a_1, \theta)$ for all (a_1, θ);

$$\beta_2^*(f, \theta, a_1) = \begin{cases} g_{(f,\theta,a_1)}, \text{ if} & \begin{aligned} u_2(a_1, b_{2,g_{(f,\theta,a_1)}}(a_1, \theta), \theta) - g_{(f,\theta,a_1)}(a_1, b_{2,g_{(f,\theta,a_1)}}(a_1, \theta)) \\ > u_2(a_1, b_{2,f}(a_1, \theta), \theta) - f(a_1, b_{2,f}(a_1, \theta)) \end{aligned} \\ b_{2,f}(a_1, \theta), & \text{otherwise} \end{cases}$$

for any $f \neq f^*$ and (θ, a_1); $\beta_2^*(f, \theta, a_1, g, y) = b_{2,g}(a_1, \theta)$ and $\beta_2(f, \theta, a_1, g, n) = b_{2,f}(a_1, \theta)$ for all $f \neq f^*$

and (a_1,θ,g); $\beta_2(f^*,\theta,a_1,g,n) = b_2^*(a_1,\theta)$ for all (a_1,θ,g);

$$\beta_3^*(I_3(f^*,a_1,g)) = \begin{cases} y, & g(a_1,b_{2,g}(a_1,\theta)) > f^*(a_1,b_2^*(a_1,\theta)) \quad \forall\theta \\ n, & \text{otherwise} \end{cases}$$

and

$$\beta_3^*(I_3(f,a_1,g)) = \begin{cases} y, & \text{if } g(a_1,b_{2,g}(a_1,\theta)) \geq f(a_1,b_{2,f}(a_1,\theta)) \quad \forall\theta \\ n, & \text{otherwise} \end{cases}$$

for any a_1, g and $f \neq f^*$; $\mu^*(\mathscr{C})(f^*) = 1$; $\mu^*(I_3(f^*,a_1,g))(\theta) = p(\theta)$ if $g(a_1,b_{2,g}(a_1,\theta)) > f^*(a_1,b_2^*(a_1,\theta))$ for all θ and $\mu^*(I_3(f^*,a_1,g))(\theta') = 1$ if there exists θ' such that $f^*(a_1,b_2^*(a_1,\theta')) \geq g(a_1,b_{2,g}(a_1,\theta'))$; For any $f \neq f^*$ and (a_1,g), $\mu^*(I_3(f,a_1,g))(\theta) = p(\theta)$ if $g(a_1,b_{2,g}(a_1,\theta)) \geq f(a_1,b_{2,f}(a_1,\theta))$ for all θ and $\mu^*(I_3(f,a_1,g))(\theta') = 1$ if there exists θ' such that $f(a_1,b_{2,f}(a_1,\theta')) > g(a_1,b_{2,g}(a_1,\theta'))$. This assessment induces (b_1^*,b_2^*) and is a renegotiation-proof perfect Bayesian equilibrium.

[Only if] Suppose that $\Gamma_R(G)$ has a renegotiation-proof perfect Bayesian equilibrium (β^*,μ^*) that induces (b_1^*,b_2^*). Letting $\beta_2^*(\emptyset) = f^*$, we have $\beta_1^*(\mathscr{C}) = b_1^*$, $\beta_2(f^*,\theta,a_1) = b_2^*(a_1,\theta)$ for all (a_1,θ), and $\mu^*(\mathscr{C})(f^*) = 1$. Sequential rationality of player 1 implies that

$$b_1^* \in \operatorname*{argmax}_{a_1} U_1(a_1,b_2^*) \tag{25}$$

whereas that of player 2 implies $u_2(a_1,b_2^*(a_1,\theta),\theta) - f^*(a_1,b_2^*(a_1,\theta)) \geq u_2(a_1,b_2^*(a_1,\theta'),\theta) - f^*(a_1,b_2^*(a_1,\theta'))$ for all a_1 and θ,θ', which, together with increasing differences, implies that b_2^* is increasing.

We also claim that

$$b_2^*(b_1^*,\theta) \in \operatorname*{argmax}_{a_2} u_2(b_1^*,a_2,\theta) \quad \forall\theta. \tag{26}$$

Suppose, for contradiction, that this is not the case for θ' and let $\hat{a}_2 \in \operatorname{argmax}_{a_2} u_2(b_1^*,a_2,\theta')$ and define $\varepsilon = u_2(b_1^*,\hat{a}_2,\theta') - u_2(b_1^*,b_2^*(b_1^*,\theta'),\theta') > 0$. Define $f'(a_1,a_2) = F^*(b_1^*,b_2^*) + \varepsilon/2$ and note that the third-party accepts f'. Assume first that f' is not renegotiated after b_1^* and note that sequential rationality of player 2 implies that $\beta_2^*(f',\theta,b_1^*) \in \operatorname{argmax}_{a_2} u_2(b_1^*,a_2,\theta)$. Let $b_{2,f'}(a_1,\theta) = \beta_2^*(f',\theta,a_1)$. Player 2's expected payoff under f' is

$$U_2(b_1^*,b_{2,f'}) - F^*(b_1^*,b_2^*) - \varepsilon/2 > U_2(b_1^*,b_2^*) - F^*(b_1^*,b_2^*)$$

contradicting that (β^*,μ^*) is a PBE. A similar argument goes through if f' is renegotiated after b_1^*.

Therefore, by (25) and (26), (b_1^*,b_2^*) is a Bayesian Nash equilibrium of G and b_2^* is increasing. Finally, suppose that b_2^* is not renegotiation-proof. This implies that for any contract f such that (f,b_2^*) is incentive compatible, there exist a_1',θ', and an incentive compatible (g,b_2) such that $u_2(a_1',b_2(a_1',\theta'),\theta') - g(a_1',b_2(a_1',\theta')) > u_2(a_1',b_2^*(a_1',\theta'),\theta') - f(a_1',b_2^*(a_1',\theta'))$ and $g(a_1',b_2(a_1',\theta)) > f(a_1',b_2^*(a_1',\theta))$ for all θ. This implies that, in any perfect Bayesian equilibrium, after history (f,θ',a_1') player 2 strictly prefers to renegotiate and offer g and the third-party accepts it. In other words, there exists no renegotiation-proof perfect Bayesian equilibrium which induces (b_1^*,b_2^*), completing the proof. □

Proof of Corollary 3. Given Corollary 2, we only need to prove that $\underline{U}_1^R = \max_{a_1} U_1(a_1,b_2^R)$. We first need the following definition

Definition 11. For any $b_2 \in A_2^{A_1 \times \Theta}$ we say that (a_1, i), $i \in \{1, 2, \ldots, n\}$ has *right deviation* (*left deviation*) at b_2 if there exists an $a_2 \in A_2$ such that $a_2 \succsim_2 b_2(a_1, \theta^i)$ $(b_2(a_1, \theta^i) \succsim_2 a_2)$ and $u_2(a_1, a_2, \theta^i) > u_2(a_1, b_2(a_1, \theta^i)\theta^i)$. Otherwise, we say that i has *no right deviation* (*no left deviation*)at b_2.

For any $b_2 \in A_2^{A_1 \times \Theta}$ and (a_1, i), $i \in \{1, \cdots, n\}$, that has right deviation at b_2, define

$$R(a_1, i) = \{k > i : b_2(a_1, \theta^k) \in BR_2(a_1, \theta^k) \text{ and } i < j < k \text{ implies that } (a_1, j) \text{ has no left deviation at } b_2\}$$

Similarly, for any (a_1, i) with $i \in \{1, \cdots, n\}$, that has a left deviation at b_2, define

$$L(a_1, i) = \{k < i : b_2(a_1, \theta^k) \in BR_2(a_1, \theta^k) \text{ and } k < j < i \text{ implies that } (a_1, j) \text{ has no right deviation at } b_2\},$$

We need the following lemma:

Lemma 2. *b_2^* is renegotiation-proof if for any (a_1, i_1) that has right deviation and any (a_1, i_2) that has left deviation at b_2^*, $R(a_1, i_1) \neq \emptyset$, $L(a_1, i_2) \neq \emptyset$, and $i_1 < i_2$ implies $R(a_1, i_1) \cap L(a_1, i_2) \neq \emptyset$.*

Proof of Lemma 2. Similar to the proof of Lemma 6 in Gerratana and Koçkesen (2012). ☐

We can now proceed to the proof of Corollary 3. We first prove that b_2^R is renegotiation proof. Fix a_1 and assume u_1 is increasing in a_2. Then, there is no (a_1, i) with left deviation by construction of b_2^R. For any (a_1, i) with right deviation, we have $n \in R(a_1, i)$. Similarly, if u_1 is decreasing in a_2. Lemma 2, therefore, implies that b_2^R is renegotiation-proof.

We next prove that for any a_1 and renegotiation-proof strategy $b_2 \in B_2^R$, we have $U_1(a_1, b_2) \geq U_1(a_1, b_2^R)$. We will use the following lemma

Lemma 3. *If $b_2 \in A_2^{A_1 \times \Theta}$ is renegotiation-proof, then (a_1, θ^n) has no right deviation at b_2 for any $a_1 \in A_1$*

Proof of Lemma 3. Similar to the proof of Lemma 7 in Gerratana and Koçkesen (2012). ☐

Fix a_1 and assume that u_1 is increasing in a_2. Let $b_2 \in B_2^R$. Lemma 3 implies that $b_2(a_1, \theta^n) \succsim_2 b_2^n(a_1)$ and hence $u_1(a_1, b_2(a_1, \theta^n), \theta^n) \geq u_1(a_1, b_2^n(a_1), \theta^n) = u_1(a_1, b_2^R(a_1, \theta^n), \theta^n)$. Also, $u_1(a_1, b_2(a_1, \theta), \theta) \geq u_1(a_1, \underline{a}_2, \theta)$ for all θ, which implies that $U_1(a_1, b_2) \geq U_1(a_1, b_2^R)$. Therefore, $\underline{U}_1^R = \max_{a_1} U_1(a_1, b_2^R)$. ☐

References

[1] Beaudry, P. and M. Poitevin (1995), "Contract Renegotiation: A Simple Framework and Implications for Organization Theory," *Canadian Journal of Economics*, 28, 302-335.

[2] Bensaid, B. and R. J. Gary-Bobo (1993), "Commitment Value of Contracts under Renegotiation Constraints," *Econometrica*, 61, 1423-29.

[3] Bhaskar, V. (2009), "Games Played in a Contracting Environment," *Games and Economic Behavior*, 67, 389-398.

[4] Bolton, P. and M. Dewtripont (2005), *Contract Theory*, Cambridge: The MIT Press.

[5] Bolton, P. and D. S. Scharfstein (1990), "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80, 93-106.

[6] Brander, J. A. and T. R. Lewis (1986), "Oligopoly and Financial Structure: The Limited Liability Effect," *American Economic Review*, 76, 956-970.

[7] Brander, J. A. and B. J. Spencer (1985), "Export Subsidies and International Market Share Rivalry," *Journal of International Economics*, 18, 83-100.

[8] Caillaud, B., B. Jullien, and P. Picard (1995), "Competing Vertical Structures: Precommitment and Renegotiation," *Econometrica*, 63, 621-646.

[9] Cho, I. K. and D. Kreps (1987), "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102, 179-221.

[10] Dewatripont, M. (1988), "Commitment through Renegotiation-Proof Contracts with Third Parties," *Review of Economic Studies*, 55, 377-389.

[11] Eaton, J. and G. M. Grossman (1986), "Optimal Trade and Industrial Policy under Oligopoly," *Quarterly Journal of Economics*, 101, 383-406.

[12] Fershtman, C. and K. Judd (1987), "Equilibrium Incentives in Oligopoly," *American Economic Review*, 77, 927-940.

[13] Fershtman, C., K. Judd, and E. Kalai (1991), "Observable Contracts: Strategic Delegation and Cooperation," *International Economic Review*, 32, 551-559.

[14] Fudenberg, D. and J. Tirole (1991), "Perfect Bayesian and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236-260.

[15] Gerratana E. and L. Koçkesen (2012), "Strategic Effects of Renegotiation-Proof Contracts," *The B.E. Journal of Theoretical Economics (Advances)*, 12:1.

[16] Holmstrom, B. and R. Myerson (1983), "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51, 1799-1819.

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

[18] Katz, M. (1991), "Game-Playing Agents: Unobservable Contracts as Precommitments," *Rand Journal of Economics*, 22, 307-328.

[19] Katz, M. (2006), "Observable Contracts as Commitments: Interdependent Contracts and Moral Hazard," *Journal of Economics & Management Strategy*, 15, 685-706.

[20] Koçkesen, L. (2007), "Unobservable Contracts as Precommitments," *Economic Theory*, 31, 539-552.

[21] Koçkesen, L. and E. A. Ok (2004), "Strategic Delegation by Unobservable Incentive Contracts," *Review of Economic Studies*, 71, 397-424.

[22] Koçkesen, L., E. A. Ok, and R. Sethi (2000), "The Strategic Advantage of Negatively Interdependent Preferences," *Journal of Economic Theory*, 92, 274-299.

[23] Mangasarian, O. L. (1994) *Nonlinear Programming*, New York: McGraw-Hill.

[24] Maskin, E. and J. Tirole (1992),"The Principal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 60, 1-42.

[25] Polo, M. and P. Tedeschi (2000), "Delegation Games and Side-Contracting," *Research in Economics*, 54, 101-116.

[26] Prat, A. and A. Rustichini (2003), "Games Played through Agents," *Econometrica*, 71, 989-1026.

[27] Sklivas, S. D. (1987), "The Strategic Choice of Managerial Incentives," *Rand Journal of Economics*, 18, 452-458.

[28] Snyder, C. M. (1996), "Negotiation and Renegotiation of Optimal Financial Contracts Under the Threat of Predation," *Journal of Industrial Economics*, 44, 325-343.

[29] Spencer, B. J. and J. A. Brander (1983), "International R & D Rivalry and Industrial Strategy," *Review of Economic Studies*, 50, 707-722.

[30] Vickers, J. (1985), "Delegation and the Theory of the Firm," *Economic Journal, Supplement*, 95, 138-147.

[31] Walsh, C. E. (1995), "Optimal Contracts for Central Bankers," *American Economic Review*, 85, 150-167.