Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Economic Theory 148 (2013) 473–501

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Ex post renegotiation-proof mechanism design ☆

Zvika Neeman [a,1], Gregory Pavlov [b,*]

[a] The Eitan Berglas School of Economics, Tel-Aviv University, Tel-Aviv, 69978, Israel

[b] Department of Economics, University of Western Ontario, Social Science Centre, London, Ontario N6A 5C2, Canada

Received 11 February 2010; final version received 10 August 2012; accepted 20 August 2012

Available online 3 January 2013

Abstract

We study what kind of equilibria of which mechanisms are ex post renegotiation-proof (EPRP), i.e., robust against the possibility of ex post renegotiation under a variety of renegotiation procedures, and which social choice functions are EPRP implementable. In complete information environments with two agents only budget balanced Groves allocations are EPRP implementable, while with three or more agents – all ex post efficient allocations are. In environments with independent private values essentially only the budget balanced "Groves in expectations" allocations are EPRP implementable, while with three or more agents and correlated beliefs – all ex post efficient allocations are.

JEL classification: D02; D70; D82; D86

Keywords: Mechanism design; Implementation; Ex post renegotiation

☆ We would like to thank Eddie Dekel, Françoise Forges, Maria Goltsman, Johannes Hörner, Bart Lipman, Dilip Mookherjee, Michel Poitevin, Pasquale Schiraldi and the seminar participants at Arizona State University, Boston University, the Hebrew University, Université de Montréal, Tel Aviv University, Toulouse, Vienna, University of Western Ontario, CETC (Montreal, 2007), Decentralization conference (Ann Arbor, 2007), World Congress of the Game Theory Society (Evanston, 2008), and mechanism design conference (Bonn, 2009) for helpful comments and conversations. We are also grateful to the editor and three anonymous referees for their thoughtful comments. Neeman's research was supported in part by the ISF and by the Google Inter-university center for Electronic Markets and Auctions. All remaining errors are ours.

* Corresponding author. Fax: +1 519 661 3666.

E-mail addresses: zvika@post.tau.ac.il (Z. Neeman), gpavlov@uwo.ca (G. Pavlov).

URLs: http://zvikaneeman.weebly.com/ (Z. Neeman), http://economics.uwo.ca/faculty/pavlov/ (G. Pavlov).

[1] Fax: +972 3 6409908.

0022-0531/$ – see front matter

http://dx.doi.org/10.1016/j.jet.2012.08.003

1. Introduction

One of the practical concerns of mechanism design theory is that players might often have incentives to change the rules of the game they are playing. Although in some cases the mechanism designer can prevent such changes, in many situations it is impossible or nearly impossible to do so, especially when a change in the rules of the game, contract, or mechanism is mutually beneficial for the agents. Such mutually consensual changes, which are known as *renegotiation*, can occur at different stages of the contractual process. *Interim renegotiation* takes place before the mechanism is played and involves a change of the mechanism and the equilibrium the players intend to play. *Ex post renegotiation* takes place after the mechanism is played and involves a change of the outcome or recommendation proposed by the mechanism. The consequences of both interim and ex post renegotiation crucially depend on the details of the renegotiation process: what alternative outcomes or mechanisms are considered? How do the players communicate with each other, and how do they select among the alternative proposals? How is the surplus that is generated by renegotiation shared among the players? Etc.

This paper is devoted to the subject of ex post renegotiation.[2] We study the problem of a mechanism designer who is unable to prevent renegotiation and is ignorant of the exact way in which renegotiation will proceed. More specifically, we are interested in the question of what kind of equilibria of which mechanisms are *renegotiation-proof*, that is, can be expected to be "stable" in the sense of surviving in their original form under a variety of different renegotiation procedures. We also want to know which social choice functions can be implemented in a way that is renegotiation-proof.

The prospect of ex post renegotiation can undermine an equilibrium of a mechanism in two different ways. First, the equilibrium outcome that is reached by the mechanism can be renegotiated. Second, the incentive compatibility of the original equilibrium may be compromised because some agent may find it in his interest to deviate from the original equilibrium when the mechanism is played in anticipation of subsequent renegotiation. This second possibility can be illustrated by the following example. Garratt and Tröger [20] consider a single object first price auction with two bidders, one that has a positive valuation for the object and one that values the object at zero. They show that if resale is not allowed, then there is a unique equilibrium in which both bidders bid zero and the high value bidder always wins the object. But if resale (or renegotiation) is allowed, then the above equilibrium is no longer incentive compatible, because the second bidder could deviate to a positive bid and try to resell the good to the first bidder. In the game with resale there is a unique equilibrium with resale on the equilibrium path. The former equilibrium is efficient, but the latter is not.

The standard approach to renegotiation that is used in the literature assumes that the mechanism designer can anticipate exactly how any outcome will be renegotiated. The anticipated renegotiation can then be incorporated into the original contract, and (in many cases) without loss of generality attention can then be limited to renegotiation-proof mechanisms.[3] In contrast, our notion of renegotiation-proofness requires that an equilibrium of a mechanism survives under all plausible renegotiation procedures and involves a considerable loss of generality,

[2] An extension of our approach to interim renegotiation is available in the working paper version of this paper. Ex post renegotiation generally imposes more constraints on the attainable outcomes than interim renegotiation. Segal and Whinston [35] and Watson [40] observe that this is the case in complete information environments, and Beaudry and Poitevin [7] make a similar observation in the case of incomplete information environments.

[3] This result, which is known as the "renegotiation-proofness principle", was first introduced by Dewatripont [14].

especially, but not only, in cases where renegotiation-proof mechanisms fail to exist as shown below.[4]

We consider quasi-linear environments with any number of agents under both complete and incomplete information. After presenting the basic setup in Section 2, in Section 3 we study ex post renegotiation in complete information environments. We show that ex post renegotiation-proof equilibria always result in ex post efficient outcomes (Lemma 1), and that for the case of three or more agents every ex post efficient social choice function can be implemented in a way that is ex post renegotiation-proof by a simple mechanism that requires two agents to report the state of the world and requires them to pay high penalties to other agents if they disagree (Proposition 2). When there are only two agents it is impossible to penalize both agents simultaneously when they disagree about the state of the world because such penalties would be undone by renegotiation. In this case, we show that the set of implementable social choice functions is strictly smaller than in the case of three or more agents: all ex post renegotiation-proof implementable social choice functions are outcome equivalent to budget balanced Groves mechanisms (Proposition 1). Many mechanism design environments do not admit the existence of budget balanced Groves mechanisms. In such cases, it is impossible to implement any social choice function in a way that is ex post renegotiation-proof. The implications and limitations of this result are discussed in Section 3.3.

Modeling ex post renegotiation in incomplete information environments is a more challenging task than in the case of complete information. In Section 4 we develop one possible approach under which the agents employ an incentive compatible renegotiation function through which they share their private information and decide collectively on how to renegotiate the outcome. If agents' values are interdependent then there may exist ex post inefficient renegotiation-proof equilibria (Example 5), but in the case of private values ex post renegotiation-proof equilibria are always ex post efficient (Lemma 2). When there are three or more agents and the agents' private information is correlated, we show that every ex post efficient social choice function can be implemented in a way that is ex post renegotiation-proof (Proposition 4). For the case of independent private values we show that every budget balanced "Groves in expectations" social choice function is ex post renegotiation-proof implementable.[5] The opposite direction requires that no residual uncertainty remains after the mechanism is played (Proposition 3). As before, some mechanism design environments may not permit the implementation of any social choice function in a way that is ex post renegotiation-proof. The implications and limitations of this result in the context of incomplete information environments are discussed in Section 4.4.

All proofs are relegated to Appendix A unless otherwise stated.

2. Setup

A group of n agents must reach an agreement that involves the choice of a *social alternative* a from a set A and *monetary transfers* to the agents, $t = (t_1, \ldots, t_n)$. An *outcome*

[4] There is a small literature that studies ex post renegotiation under multiple possible renegotiation procedures. In environments with complete information Sjöström [37] and Amoros [2] study the case of three or more agents and Rubinstein and Wolinsky [34] study a specific buyer and seller problem. Forges [17,18] studies ex post renegotiation under incomplete information; Ausubel and Cramton [5] allow for multiple resale scenarios following the Vickrey auction.

[5] "Groves in expectations" mechanisms are equivalent to Groves mechanisms from the agents' interim perspective. See, for example, Williams [41].

(a, t) of the process of negotiation among the agents is said to be *feasible* if $a \in A$ and $t \in \mathbb{T} = \{t \in \mathbb{R}^n\colon \sum_{i=1}^n t_i \leqslant 0\}$.[6]

The agents' preferences over outcomes depend on the state of the world θ that is chosen from a finite set Θ.[7] Each agent i is an expected utility maximizer with a quasi-linear payoff function $v_i(a, \theta) + t_i$, where $v_i : A \times \Theta \to \mathbb{R}$ describes his preferences over social alternatives for different states of the world, and t_i denotes the monetary transfer he receives. We assume that for every state θ there is a single alternative $a^*(\theta)$ that maximizes the total surplus $\sum_{i=1}^n v_i(a, \theta)$; this alternative is called *ex post efficient in state* θ.[8] A vector of transfers t that satisfies $\sum_{i=1}^n t_i = 0$ is called *budget balanced*. An outcome (a, t) is called *ex post efficient in state* θ if $a = a^*(\theta)$ and t is budget balanced.

A *social choice function* is a mapping $f : \Theta \to A \times \mathbb{T}$ from the set of states into feasible outcomes. A social choice function f is said to be *ex post efficient* if for every $\theta \in \Theta$ the outcome $f(\theta) = (a(\theta), t(\theta))$ is ex post efficient in every state θ.

We distinguish between complete and incomplete information environments as follows:

Complete information. Complete information environments are analyzed in Section 3. The state of the world θ is commonly known among the agents. It is convenient to denote the set of possible preferences of agent i by Θ_i. With this notation, the state of the world can be represented by a vector of the agents' payoff relevant types: $\theta = (\theta_1, \ldots, \theta_n)$, where $\theta_i \in \Theta_i$, and, with a slight abuse of notation, $v_i(a, \theta) = v_i(a, \theta_i)$ for every i and a.

While in general $\Theta \subseteq \prod_{i=1}^n \Theta_i$, for some results we assume that the set of states has a structure of a Cartesian product.[9]

Condition 1 *(Product domain).* $\Theta = \prod_{i=1}^n \Theta_i$.

A *mechanism* $\langle S, m\rangle$ specifies a set of messages S_i for each agent i and an outcome rule $m : S \to \Delta(A \times \mathbb{T})$ from the set of message profiles $S = \prod_{i=1}^n S_i$ into the set of lotteries over feasible outcomes. A mechanism $\langle S, m\rangle$ together with a state of the world θ defines a complete information game, where a strategy of agent i is $\sigma_i(\theta) \in \Delta(S_i)$. We say that $\sigma = (\sigma_1, \ldots, \sigma_n)$ is an *equilibrium* if for every $\theta \in \Theta$ the strategy profile $(\sigma_1(\theta), \ldots, \sigma_n(\theta))$ is a Nash equilibrium of the complete information game in state θ. Equilibrium σ is said to be *ex post efficient* if the equilibrium outcome is ex post efficient with probability one in every state $\theta \in \Theta$.

A social choice function f is said to be *implementable* if there exists a mechanism $\langle S, m\rangle$ that has an equilibrium σ such that the equilibrium outcome coincides with $f(\theta)$ in every state θ.[10]

Incomplete information. Incomplete information environments are analyzed in Section 4. The set of states of the world is $\Theta = \prod_{i=1}^n \Theta_i$, there is a common prior distribution P over Θ, and

6 In some environments the right-hand side of the feasibility constraint on the sum of monetary transfers may depend on the chosen social alternative. Using a standard argument it is possible to redefine the agents' payoffs in a way that is consistent with our model. See, for example, Fudenberg and Tirole [19, Chapter 7].

7 We assume that Θ is finite to simplify the exposition, but our results continue to hold if Θ is infinite.

8 This assumption is generically satisfied if Θ is finite.

9 This is a standard condition in implementation theory (sometimes called "independent domain"). See, for example, Moore [31].

10 Thus we employ a weak notion of implementation.

each agent i privately observes his type $\theta_i \in \Theta_i$. We consider both private and interdependent values environments; by private values we mean the following.

Condition 2 *(Private values). The payoff of each agent depends only on his own type, i.e., with a slight abuse of notation*:

$$v_i(a, \theta_i, \theta_{-i}) = v_i(a, \theta_i) \quad \forall i \in \{1, \dots, n\},\ a \in A,\ \theta_i \in \Theta_i,\ \theta_{-i} \in \Theta_{-i}.$$

Like in the case of complete information, a *mechanism* $\langle S, m\rangle$ specifies a set of messages S_i for each agent i and an outcome rule $m : S \to \Delta(A \times \mathbb{T})$ from the set of message profiles $S = \prod_{i=1}^n S_i$ into the set of lotteries over feasible outcomes. A mechanism $\langle S, m\rangle$ together with a common prior P defines a Bayesian game, where a strategy of agent i is $\sigma_i : \Theta_i \to \Delta(S_i)$. We say that $\sigma = (\sigma_1, \dots, \sigma_n)$ is an *equilibrium* if it is a Bayesian Nash equilibrium of the Bayesian game. Equilibrium σ is said to be *ex post efficient* if the equilibrium outcome is ex post efficient with probability one in every state $\theta \in \Theta$.

A social choice function f is said to be *implementable* if there exists a mechanism $\langle S, m\rangle$ that has an equilibrium σ such that the equilibrium outcome is payoff equivalent to f from the agents' interim perspective. That is, the equilibrium expected payoff of agent i of type $\theta_i \in \Theta_i$ is equal to

$$E_{\theta_{-i}|\theta_i}\big[v_i\big(a(\theta), \theta\big) + t_i(\theta)\big] = \sum_{\theta_{-i} \in \Theta_{-i}} P(\theta_{-i} \mid \theta_i)\big(v_i\big(a(\theta), \theta\big) + t_i(\theta)\big)$$

where $(a(\theta), t(\theta)) = f(\theta)$ for every $\theta \in \Theta$.

Although we do not discuss the subject of individual rationality in this paper (except for Section 4.4), it is straightforward to incorporate into the analysis an appropriate notion of individual rationality as an additional set of constraints on the social choice function and mechanism.

3. Ex post renegotiation-proofness under complete information

3.1. Preliminaries

In a complete information environment a state of the world $\theta \in \Theta$ is realized and becomes commonly known among the agents, but is not known to the mechanism designer. The agents then play the game induced by a given mechanism $\langle S, m\rangle$. After the outcome of playing the mechanism is determined, the agents have an opportunity to change it to some other outcome according to some renegotiation procedure.[11] Our goal is to determine which equilibria of which mechanisms are robust against such renegotiation.

Rather than modeling various renegotiation procedures as games, we take the following shortcut. Suppose an outcome (a, t) was reached by the play of the mechanism in state θ. We say that the outcome (a, t) *can be renegotiated* if there exists another feasible outcome (a', t') that Pareto dominates (a, t) in state θ. That is, in this state of the world all the agents weakly prefer, and at

[11] Alternatively one could allow renegotiation to take place before the lotteries from $\Delta(A \times \mathbb{T})$ prescribed by the outcome rule m are carried through. Since the agents have quasi-linear payoffs such an approach would not affect the results, but it would make the arguments more cumbersome.

least one agent strictly prefers, the alternative outcome (a', t') to the original outcome (a, t) that was prescribed by the mechanism.[12]

Consider an equilibrium of some mechanism. If we allow the possibility of ex post renegotiation, then an equilibrium outcome may be undermined in two ways. First, the equilibrium outcome reached by the mechanism may be renegotiated. Second, the original equilibrium strategy profile may cease to be an equilibrium once ex post renegotiation is allowed, since deviating when playing the mechanism may bring about an outcome which may itself be beneficially renegotiated. These two possibilities lead to the following definition.

Definition 1. An equilibrium σ of a mechanism $\langle S, m\rangle$ is *ex post renegotiation-proof* (*EPRP*) if both of the following conditions hold:

(i) An outcome that is obtained under the equilibrium play of the mechanism cannot be renegotiated.
(ii) No agent can improve upon his equilibrium payoff in any state by a unilateral deviation from σ followed by renegotiation of the resulting outcomes.

We are also interested in social choice functions that can be implemented in a renegotiation-proof way.

Definition 2. A social choice function is *ex post renegotiation-proof* (*EPRP*) *implementable* if there exists a mechanism $\langle S, m\rangle$ that has an EPRP equilibrium σ such that the equilibrium outcome coincides with $f(\theta)$ in every state θ.

3.2. Results

In this section we study what kind of equilibria of which mechanisms are EPRP, and characterize the social choice functions that are EPRP implementable. The first requirement for an equilibrium of some mechanism to be EPRP is that the equilibrium outcomes are not Pareto dominated in any state. This is a simple implication of part (i) of Definition 1 and is stated without proof.

Lemma 1. *In a complete information environment every EPRP equilibrium is ex post efficient.*

Next we present an example where an equilibrium fails to be EPRP in spite of being ex post efficient and in dominant strategies. In particular, the example shows that ex post renegotiation-proofness not only requires the transfers to be budget balanced in every state of the world, but also imposes restrictions on the relative size of the transfers across different states of the world.

Example 1. A buyer and a seller can trade a single good. The buyer values the good at V that can be either 0 or 2, the seller values the good at 1. The realization of V and the seller's valuation are

[12] A standard approach to modeling renegotiation, like in Maskin and Moore [29], assumes that ex post renegotiation proceeds according to some given process (or a game). This process induces a mapping from outcomes and states into (possibly different) outcomes, and this mapping is commonly known by the agents and the mechanism designer. In contrast, we allow more flexibility in the ways renegotiation may proceed, and assume that the mechanism designer is ignorant of the exact way in which renegotiation will proceed. See Section 3.3 for further discussion.

commonly known between the agents. Consider a mechanism where the buyer is asked to report his value: after a report "$V = 2$" the good is transferred from the seller to the buyer at a price p_2, and after a report "$V = 0$" there is no trade and the buyer pays p_0 to the seller. It is easy to see that the buyer has a dominant strategy to report his true valuation if $p_2 - p_0 \in (0, 2)$, and the resulting outcome is ex post efficient. However, as we show below, this equilibrium is not EPRP unless $p_2 - p_0 = 1$.

Suppose $p_2 - p_0 \in (1, 2)$. If the buyer with $V = 2$ reports "$V = 0$" then the payoffs of the buyer and the seller (without renegotiation) would be $-p_0$ and p_0, respectively. This outcome is Pareto dominated by a decision to trade at a new price $\hat{p}$ that satisfies $\hat{p} - p_0 \in (1, 2)$. Hence, for any such $\hat{p} < p_2$, the buyer would prefer to misreport and then renegotiate the outcome to trade at the price $\hat{p}$ rather than report his true valuation. Thus, the original equilibrium is not EPRP.[13]

This example deals with just one simple class of mechanisms. More generally, one could ask both the buyer and the seller about the state of the world, and then condition the outcomes on their reports. However, as we argue below, allowing for more complicated mechanisms does not expand the range of implementable outcomes in the environment considered in this example.

Note that in this example we can restate the EPRP restrictions on the transfers as a requirement that the seller's payoff (without renegotiation) be independent of the report of the buyer. If it is not the case, then the buyer could make a report that minimizes the seller's status quo payoff, and renegotiate the outcome while keeping all renegotiation surplus to himself. But if the seller's payoff is a constant independent of the state of the world, then budget balance implies that the buyer is a residual claimant of the total surplus. Hence, in the EPRP equilibrium of the considered mechanism the buyer's transfer equals the externality that his report imposes on the seller.

To generalize this idea we consider a class of *Groves mechanisms* from the mechanism design literature (Groves [23]). Each agent i reports his own payoff relevant type $\theta_i \in \Theta_i$, an ex post efficient alternative given the reports is implemented, and each agent is given a transfer (as a function of the reports) that equals the externality he imposes on the other agents. There exists an equilibrium of this mechanism in dominant strategies where each agent tells the truth, and thus Groves mechanisms can be used to implement the following class of social choice functions.

Definition 3. Suppose there are two agents. A social choice function f is called *Groves* if $f(\theta) = (a^*(\theta), t(\theta))$ where a^* is ex post efficient and transfers t satisfy

$$t_i(\theta_i, \theta_j) = v_j\big(a^*(\theta_i, \theta_j), \theta_j\big) + H_i(\theta_j) \quad \forall i \in \{1, 2\},\ (\theta_i, \theta_j) \in \Theta, \tag{1}$$

for some function $H_i : \Theta_j \to \mathbb{R}$.

Next we describe which social choice functions for the case of two agents can be implemented in a renegotiation-proof way. If the *product domain* restriction (Condition 1) is satisfied, then only Groves social choice functions that are budget balanced are EPRP implementable. Otherwise, there may be other EPRP implementable social choice functions as well.[14]

[13] The argument for the case $p_2 - p_0 \in (0, 1)$ is similar. The buyer with $V = 0$ will find it profitable to report "$V = 2$" and then renegotiate to "no trade".

[14] See Example 2 below.

Proposition 1. *Consider a complete information environment with two agents.*

(i) *Every budget balanced Groves social choice function is EPRP implementable.*
(ii) *Suppose the product domain condition is satisfied, and a social choice function f is EPRP implementable. Then f is budget balanced Groves.*

To prove part (i) of the result we consider a budget balanced Groves mechanism that corresponds to a given budget balanced Groves social choice function. This mechanism has an ex post efficient equilibrium where each agent truthfully reports his payoff relevant type. Moreover, like in Example 1, the truthful equilibrium of this mechanism has a property that each agent's payoff (without renegotiation) is independent of the report of the opponent.[15] Using the terminology of Segal and Whinston [36] in this case there are "no harmful direct externalities" in the absence of renegotiation. Hence, any strategy of misreporting one's own type and subsequent renegotiation is unprofitable: it can neither lower the opponent's payoff, nor increase the total surplus (since the equilibrium outcome is ex post efficient).[16]

To illustrate the idea behind the proof of part (ii) of the result, first note that by Lemma 1 only ex post efficient social choice functions can potentially be EPRP implementable. Suppose we want to implement some ex post efficient social choice function that is not budget balanced Groves. A simple mechanism that works when there is no renegotiation is as follows. Each agent is asked to report the whole state of the world (and not just his payoff relevant type as in Groves mechanisms). If the agents' reports agree, the outcome prescribed by the social choice function is implemented; if they disagree, both agents are penalized. However, penalties on both agents must involve ex post inefficient outcomes, which can be undone by renegotiation. Hence, if we want to sustain truth-telling in such a direct revelation mechanism, we need to ensure that no agent can profit from misreporting the state and renegotiating the resulting outcome.

Consider a simple case where agent 1 has two possible types, θ_1 and θ_1', and the preferences of agent 2 are fixed. Denote by $U_2(\alpha, \beta)$ the payoff of agent 2 in a direct revelation mechanism (without renegotiation) when agent 1's report is α and agent 2's report is β, where $\alpha, \beta \in \{\theta_1, \theta_1'\}$. The truth-telling conditions for agent 2 are

$$U_2(\theta_1, \theta_1) \geqslant U_2(\theta_1, \theta_1') \quad \text{and} \quad U_2(\theta_1', \theta_1') \geqslant U_2(\theta_1', \theta_1). \tag{2}$$

On the other hand, it can be shown that the renegotiation-proofness constraints for agent 1 boil down to the requirement of "no harmful direct externalities" of agent 1's report on agent 2's status quo payoff:

$$U_2(\theta_1', \theta_1) \geqslant U_2(\theta_1, \theta_1) \quad \text{and} \quad U_2(\theta_1, \theta_1') \geqslant U_2(\theta_1', \theta_1'). \tag{3}$$

Combining (2) and (3) we find that the payoff of agent 2 must be independent of the reports. Hence, agent 1 must be a residual claimant of the total surplus as in a Groves mechanism. This

[15] Suppose agent i has type θ_i and reports it truthfully, while agent j reports θ_j. Then i's payoff is

$$v_i\big(a^*(\theta_i, \theta_j), \theta_i\big) + t_i(\theta_i, \theta_j) = v_i\big(a^*(\theta_i, \theta_j), \theta_i\big) - t_j(\theta_i, \theta_j) = -H_j(\theta_i)$$

where the first equality is by budget balance, and the second by (1).

[16] Segal and Whinston [36] consider an incomplete contracts model where the parties make ex ante investments that are noncontractible. In Proposition 5 they show that first best investments can be achieved if there are "no harmful direct externalities" between the players, and ex post bargaining is efficient. We thank a referee for bringing a connection with this result to our attention.

argument implies that in the environment considered in Example 1 only budget balanced Groves social choice functions are EPRP implementable.

If only one agent's payoff depends on the state of the world, then we can repeat the argument above for every pair of states, and show that every EPRP implementable social choice function must be budget balanced Groves.[17] If the preferences of both agents depend on the state of the world, then the argument above can be used to link every pair of states where the preferences of only one agent vary. For product domain environments this generates enough restrictions on EPRP implementable social choice functions to imply that they must be budget balanced Groves. If the product domain condition is not satisfied, then, as we show in the next example, there may be other social choice functions that are renegotiation-proof.

Example 2. A buyer and a seller can trade a single good. The buyer values the good at V, which can be either 0 or 2, and the seller values the good at C, which can be either 1 or 3. The realization of (V, C) is commonly known between the agents. Every budget balanced Groves social choice function prescribes trade if and only if the state is $(2, 1)$, the buyer's transfers satisfy $t_b(2, 1) - t_b(0, 1) = -1$ and $t_b(0, 3) - t_b(2, 3) = 0$, and the seller's transfers satisfy $t_s(2, 1) - t_s(2, 3) = 2$ and $t_s(0, 3) - t_s(0, 1) = 0$. In addition, budget balance requires that $t_b(V, C) + t_s(V, C) = 0$ for every (V, C). This implies that

$$\begin{aligned} 0 &= \big(t_b(2, 1) + t_s(2, 1)\big) + \big(t_b(0, 3) + t_s(0, 3)\big) \\ &\quad - \big(t_b(0, 1) + t_s(0, 1)\big) - \big(t_b(2, 3) + t_s(2, 3)\big) = 1. \end{aligned}$$

Hence, no budget balanced Groves social choice functions exist in this environment.

By part (ii) of Proposition 1 there are no EPRP implementable social choice functions if all combinations of V and C are possible. Suppose, however, that there are just two possible states: (V, C) is either equal $(0, 1)$ or $(2, 3)$. A social choice function that prescribes no trade together with a constant payment from one agent to another is EPRP implementable by a mechanism that simply ignores the agents' messages.

Next, we briefly discuss the case of three or more agents. In this case, any ex post efficient social choice function f can be EPRP implemented using the following standard construction. Agents 1 and 2 are asked to report the state of the world; if their reports agree, the outcome prescribed by f is implemented; if their reports conflict, agents 1 and 2 pay large penalties to agent 3. There exists an equilibrium where agents 1 and 2 report the state truthfully if the penalties are large enough. This equilibrium is EPRP: (i) the outcome is ex post efficient if the agents tell the truth; (ii) the penalties can be chosen to be large enough so that any strategy of sending a wrong report followed by renegotiation would be unprofitable.[18]

Proposition 2. *Consider a complete information environment with $n \geqslant 3$ agents. A social choice function f is EPRP implementable if and only if $f(\theta) = (a(\theta), t(\theta))$ is ex post efficient for every θ.*[19]

---

17 Note that the product domain condition is trivially satisfied in this case.

18 Sjöström [37] obtains a similar result in a somewhat different framework. He studies an exchange economy with agents that may be risk-neutral or risk-averse, and similarly to us assumes that the mechanism designer is ignorant of the exact way renegotiation will proceed. Theorem 1 in Sjöström [37] shows that a social choice function is uniquely implementable in undominated Nash equilibrium if and only if it is Pareto efficient.

19 The proof of Proposition 2 is available in the working paper version of the paper.

3.3. Discussion

Comparison with the standard approach. Ex post renegotiation is usually assumed to proceed according to some exogenously given process that is commonly known among the agents, and is also known to the mechanism designer who may take it into account when designing the mechanism.[20] Since many papers on renegotiation under complete information belong to the hold-up and incomplete contracts literature, let us consider one such model that was studied by Edlin and Reichelstein [15] and Che and Hausch [10].[21] A buyer and a seller make relationship-specific investments that determine their payoffs from subsequent trade. Following the investment decisions the state of the world is realized and commonly observed, and then the agents may play a game that is induced by a given mechanism. In Edlin and Reichelstein [15] the mechanism is simply a constant outcome determined by a fixed price contract, while Che and Hausch [10] allow for nontrivial revelation mechanisms. The outcomes determined by the mechanism are renegotiated so that the post-renegotiation outcome is ex post efficient, and the renegotiation surplus is shared between the agents in fixed proportions. This process of ex post renegotiation uniquely determines a mapping from the outcomes reached by the mechanism and states into (possibly other) outcomes.

Our approach to ex post renegotiation assumes that the mechanism designer is uncertain about how renegotiation works in this context. It may be unclear to him whether renegotiation always results in ex post efficient outcomes, and the sharing rule may be unknown to him and may vary depending on circumstances. In such a case, a mechanism designer who is interested in a particular social choice function may prefer to avoid uncertain renegotiation and try to achieve implementation through an equilibrium of a mechanism that is renegotiation-proof. Even if the mechanism designer is primarily interested in efficiency and not so much in how the surplus is distributed among the agents, then our approach can be useful if renegotiation is uncertain and possibly inefficient. The downside of our approach, however, is that EPRP implementable social choice functions may fail to exist, which we discuss next.

Existence of EPRP mechanisms and social choice functions. As explained above, EPRP implementable social choice functions may fail to exist when there are only two agents. If the preferences of one of the agents are independent of the state of the world, then it is easy to construct budget balanced Groves mechanisms using the same argument as in Example 1. However, as shown in Example 2, existence is not guaranteed when the preferences of both agents depend on the state of the world and the product domain condition is satisfied. To get a better sense of the scope of this nonexistence note that Groves mechanisms implement ex post efficient alternatives in dominant strategies, and are the only mechanisms that do so if each agent's set of types is a connected open subset of an Euclidean space and payoffs are continuously differentiable in types.[22] On the other hand, by part (ii) of Proposition 1, EPRP implementable social choice functions are equivalent to budget balanced Groves social choice functions, whether or not the aforementioned technical condition is satisfied. Hence, in the case of two players and product domain the set of EPRP mechanisms may be a strict subset

[20] See, for example, Maskin and Moore [29]. Some papers, like Chung [11] and Aghion et al. [1] assume that the mechanism designer has the power to affect the way the renegotiation will proceed.

[21] Segal and Whinston [35] discuss other papers that use a similar approach.

[22] See, for example, Holmström [24].

of the set of budget balanced efficient dominant strategy mechanisms, which often fail to exist.[23]

We realize that the EPRP requirements are strong, and so it is not surprising that some environments admit no EPRP mechanisms and implementable social choice functions. We hope that our theory of renegotiation-proof mechanisms can be viewed as a theory of "stable" institutions. If a mechanism retains its original form under a wide variety of circumstances, then such a mechanism must be renegotiation-proof in our strong sense. If, on the other hand, such a renegotiation-proof mechanism does not exist in a given environment, then one should not expect to observe an institution for governing transactions whose rules remain stable and not subject to perpetual renegotiation.

What next? There are several ways to weaken or circumvent the EPRP requirements.

1. The mechanism designer may be interested in implementing social choice correspondences, rather than social choice functions. For example, suppose the designer cares only about efficiency and not about how the surplus is distributed among the agents.[24] If renegotiation is uncertain, but we know that it leads to efficient outcomes, then there is no reason for the designer to restrict attention to renegotiation-proof mechanisms, and she should be content with any mechanism. More generally, with each given mechanism one could associate an outcome correspondence, which for every state gives the set of possible equilibrium outcomes of the play of the mechanism followed by uncertain renegotiation. The designer can then choose a mechanism by comparing the induced outcome correspondences of the mechanisms according to her preferences.[25]

2. Renegotiation can be modeled as being less permissive, or being more predictable. For example, in certain situations renegotiation is costly. In other cases exogenous restrictions may be imposed on the kinds of renegotiations that are allowed. For many specific applications the scope of renegotiation can be significantly reduced through careful contract design that pays close attention to the details of trade and renegotiation technology. For example, Watson [40] shows that the effect of ex post renegotiation may be significantly reduced when the actions that consummate trade are modeled as inalienable actions of some agents and trade opportunities are nondurable. In related work, Evans [16] shows that ex post renegotiation loses much of its power if sending a message in a mechanism entails an individual cost for the agent who sends it.

3. The mechanism designer may try to extract information about anticipated renegotiation from the agents through the course of play of the mechanism, and to condition the outcomes of the mechanism on this information. Such mechanisms are likely to be quite complex because they must provide the agents with the incentives to reveal all their private information – their

23 See, for example, Green and Laffont [21] and Walker [39].

24 In the hold-up and incomplete contracts models it seems natural to study social choice functions, because the distribution of the surplus between the agents in different states of the world determines the incentives for ex ante investments. For example, in the case of "selfish" investments (when investment affects only an agent's own payoff) Rogerson [33] shows that mechanisms that are "Groves in expectations" (see Section 4.2) provide first-best incentives for ex ante investment. Hence, first-best investment incentives can be provided by EPRP mechanisms if there are three or more agents, and whenever budget balanced Groves mechanisms exist if there are only two agents.

25 This point mirrors an approach that has been taken by Yamashita [42] and Börgers and Smith [9] in a different context. They consider a model in which players may play any (weakly) undominated strategy, construct the induced outcome correspondences for different mechanisms, and compare them. We thank a referee for bringing the connection with this literature to our attention.

payoff relevant private information and their beliefs about the expected renegotiation scenario.[26] In contrast, we study a simpler class of mechanisms that do not attempt to extract any information about expected renegotiation from the agents, but nonetheless implement the desired social choice function. Studying such mechanisms is a natural first step towards understanding of how to do mechanism design in the presence of multiple renegotiation scenarios.[27]

4. In some two agent environments it may be possible to expand the set of EPRP implementable social choice functions by introducing a third party. This third party need not be able to observe the state of the world, its only role would be to collect the penalties from agents 1 and 2 when their reports about the state of the world disagree.[28]

4. Ex post renegotiation-proofness under incomplete information

4.1. Preliminaries

In an incomplete information environment the state of the world $\theta = (\theta_1, \ldots, \theta_n)$ is drawn according to a prior probability distribution P, each agent i learns his private information θ_i and employs Bayes' rule to update his beliefs about the other agents' types. Then, the agents play the Bayesian game that is induced by a given mechanism $\langle S, m\rangle$. After they "play the mechanism" the agents have an opportunity to change the outcome that is determined by the mechanism to some other outcome according to some renegotiation procedure. Our goal is to determine which equilibria of which mechanisms are robust against various renegotiation scenarios of this form.[29]

Suppose that the play of the mechanism produced the outcome (a, t). Under complete information we said that (a, t) can be renegotiated if there is another outcome (a', t') that Pareto dominates the original outcome given the state of the world, and thus each agent is willing to switch from (a, t) to (a', t'). There are several challenges with developing a natural counterpart of this definition for the case of incomplete information. First, the agents may still face some residual uncertainty about the state of the world even after the outcome of the mechanism is observed. The agents may have different preferences over (a, t) vs. (a', t') depending on their own private information, and depending on the private information of the other agents. Second,

[26] We are aware of only two papers that allow such complex mechanisms. Green and Laffont [22] consider a quasi-linear model with two players and costly renegotiation. They construct a renegotiation-proof mechanism for implementing the ex post efficient social choice correspondence. Amoros [2] studies unique implementation in general environments with uncertain renegotiation. His characterization results require the mechanism designer to be able to destroy the agents' endowments, which would lead to renegotiation in our model.

[27] Conceptually, the mechanisms we study relate to the complex mechanisms outlined above in a way that is similar to the way that dominant strategy (and ex post) mechanisms relate to Bayesian mechanisms. Dominant strategy (and ex post) mechanisms are "belief-free" and rely only on the agents' payoff relevant information, while Bayesian mechanisms may also rely on the agents' higher-order beliefs (Bergemann and Morris [8], Chung and Ely [12]). The simple mechanisms we study are "belief-free" with respect to anticipated renegotiation scenarios, while complex mechanisms may rely on the agents' beliefs about future renegotiation.

[28] While it is commonly argued that an introduction of third parties raises a possibility of collusion by a subset of agents, Sjöström [37] and Baliga and Sjöström [6] show that there are circumstances when this need not be the case.

[29] As in the case of complete information, the standard approach to modeling renegotiation assumes that ex post renegotiation proceeds according to some given process (or a game). This process induces a mapping from outcomes, states, and the agents' beliefs into outcomes and the agents' beliefs, and this mapping is commonly known among the agents and the mechanism designer. In contrast, we permit renegotiation to proceed in many different ways, and assume that the mechanism designer is ignorant of the exact way in which it will proceed. See Section 4.4 for further discussion.

a model of renegotiation as a collective decision on whether to switch from the status quo outcome (a,t) to a single alternative outcome (a',t') seems overly restrictive. A more plausible model of renegotiation should allow the agents to communicate with each other, reveal some additional information, and then decide to which alternative outcome to switch conditional on the revealed information.

To provide a formal definition of our renegotiation procedure we first discuss the agents' beliefs at the point immediately after they have played the mechanism but before they have started renegotiating the outcome produced by the mechanism. Fix an equilibrium strategy profile σ of a given mechanism $\langle S, m\rangle$. For every outcome (a,t) that may be potentially reached we can derive the agents' beliefs conditional on this outcome being realized. First suppose agent i of type θ_i has played his equilibrium strategy σ_i. If outcome (a,t) is on the equilibrium path, then his beliefs are derived by Bayes' rule: he takes into account the reached outcome (a,t), the equilibrium strategy profile σ, and his type θ_i. If outcome (a,t) was not supposed to happen under σ, then he may hold arbitrary beliefs.[30] Next suppose agent i of type θ_i has played σ_i' instead of his equilibrium strategy σ_i. For every outcome (a,t) the beliefs of this agent are derived by Bayes' rule: he takes into account the reached outcome (a,t), the strategy profile (σ_i', σ_{-i}), and his type θ_i.

After a given outcome (a,t) has been reached by the mechanism, we allow it to be challenged by some *renegotiation mapping* $\rho : \Theta \to \Delta(A \times \mathbb{T})$. Note that ρ can be state contingent, because we want to allow the status quo outcome to be renegotiated to different alternative outcomes depending on the state of the world. Since the information about the state of the world is dispersed among the agents, appropriate incentives must be given to the agents for them to reveal it. Specifically, a mapping $\rho : \Theta \to \Delta(A \times \mathbb{T})$ is said to be *incentive compatible* for given agents' beliefs if under these beliefs each type of each agent weakly prefers the outcome according to ρ to any other outcome according to ρ where he lies about his type and the other agents tell the truth.[31] Finally, we want the agents to be willing to switch from the outcome prescribed by the mechanism to whatever alternative outcome is generated by the renegotiation mapping ρ. Thus, ρ is called *posterior individually rational with respect to outcome* (a,t) if for every outcome (a',t') that is produced by ρ, no agent strictly prefers (a,t) to (a',t') given his beliefs and whatever has been revealed through the play of ρ.

Such an approach to ex post renegotiation appears to be quite general, and we would like to note that many alternative renegotiation procedures can be equivalently represented in this framework. For example, suppose that after a given outcome (a,t) has been reached, a single alternative outcome (a',t') is proposed, and the agents simultaneously vote whether to switch from (a,t) to (a',t'). If all agents vote in favor of switching, then (a',t') is implemented; otherwise (a,t) is implemented. The voting rules, together with given agents' beliefs, specify an ex post renegotiation game of incomplete information. For any equilibrium of this game we can construct an outcome equivalent incentive compatible renegotiation mapping ρ using a revelation

30 There exist other reasonable approaches to specifying beliefs after the outcomes that are off the equilibrium path. For example, one can view these beliefs as being part of the description of the equilibrium of the mechanism, and, since we are using a weak notion of implementation, argue that these beliefs may be "chosen" by the mechanism designer. Since none of our results depend on the way these off-equilibrium beliefs are specified we have decided to adopt a more permissive approach and allow for arbitrary beliefs.

31 Incentive compatibility can equivalently be defined as follows. Consider a direct mechanism $\rho : \Theta \to \Delta(A \times \mathbb{T})$. Together with given agents' beliefs, ρ specifies an ex post renegotiation game of incomplete information. We say that ρ is incentive compatible for given agents' beliefs if ρ admits a truthful equilibrium.

principle-like argument. Since in this voting game each agent can veto the switch to (a', t'), the equilibrium conditions guarantee that ρ satisfies posterior individual rationality with respect to (a, t).[32]

We do not claim, however, that for any outcome of any reasonable renegotiation procedure there exists an equivalent representation in our framework. Here are a few considerations. First, one may argue that after the agents have played the mechanism agent i's private information may contain other things beyond θ_i. For example, suppose the equilibrium strategy σ_i of agent i is mixed, and that the other agents are unable to infer the pure actions actually taken by agent i from observing the final outcome (a, t). In this case the private history of actions taken by agent i becomes part of i's private information. Second, in some settings the agents may receive additional signals about the state of the world after they have played the mechanism. This may relax the incentive compatibility constraints on the renegotiation mapping ρ that ensure truthful information revelation by the agents. As an extreme example, if the state of the world becomes commonly known at the ex post stage, then it is reasonable to drop the incentive compatibility constraints on ρ altogether.[33] Third, one could replace the requirements of posterior individual rationality with weaker constraints of interim individual rationality. However, we found that in many cases such an approach to renegotiation precludes the existence of any EPRP implementable social choice functions, because of the existence of profitable deviation opportunities of the following kind: suppose that agent i of type θ_i plays the equilibrium strategy of type θ_i', and, as a result, at the renegotiation stage the other agents may end up with incorrect beliefs about i. In such a case it is easy to construct renegotiation mappings that are profitable for agent i of type θ_i, and satisfy the interim individual rationality constraints of the other agents under their incorrect beliefs, but not under the true distribution over player i's types. We believe that such renegotiation scenarios are not very plausible because they fail to protect the agents from accepting renegotiation proposals that are worse for them than the status quo. Hence, it seems that some alternative approach to specifying the agents' beliefs at the point just before renegotiation is required to make sense of ex post renegotiation under interim individual rationality.

We believe that our framework can be extended to incorporate the considerations above. At the moment, however, we have chosen to proceed with our current approach. Summarizing, here is our definition of renegotiation under incomplete information.

Definition 4. Fix an equilibrium σ of a mechanism $\langle S, m\rangle$.

(i) Suppose that an outcome (a, t) can be reached with a positive probability under σ. We say that (a, t) *can be renegotiated with a positive probability on the equilibrium path* if there exists an incentive compatible renegotiation mapping ρ that is posterior individually rational with respect to (a, t), and at least one type of one agent that has a positive probability (given (a, t) and σ) strictly prefers ρ to (a, t), where the beliefs of each agent i are derived by Bayes' rule (conditional on (a, t), σ, and θ_i).
(ii) Suppose that an outcome (a, t) can be reached with a positive probability following agent j's unilateral deviation σ_j' from σ. We say that (a, t) *can be renegotiated with a positive probability off the equilibrium path* if there exists an incentive compatible renegotiation mapping ρ that is posterior individually rational with respect to (a, t), and at least one type of one

[32] We have studied this approach to renegotiation in the previous version of this paper.

[33] One of the approaches we considered in the previous version of the paper allowed for an external "oracle" device which could help the agents with renegotiation by revealing the state of the world.

agent that has positive probability (given (a,t) and (σ'_j, σ_{-j})) strictly prefers ρ, where the beliefs of agent j are derived by Bayes' rule (conditional on (a,t), (σ'_j, σ_{-j}) and θ_j), and the beliefs of each agent $i \neq j$ are derived by Bayes' rule (conditional on (a,t), σ and θ_i) whenever possible and are arbitrary otherwise.

As in the case of complete information, there are two ways in which ex post renegotiation may undermine the equilibrium of a given mechanism. First, the equilibrium outcome that is reached by the mechanism may be renegotiated. Second, the original equilibrium strategy profile may cease to be an equilibrium once ex post renegotiation is allowed, since a deviation in the mechanism would bring about an outcome which may be potentially beneficially renegotiated. These two possibilities lead to the following incomplete information analog of Definition 1 (Section 3.1).[34]

Definition 5. An equilibrium σ of a mechanism $\langle S, m\rangle$ is *ex post renegotiation-proof* (*EPRP*) if both of the following conditions hold:

(i) The outcomes obtained under the equilibrium play of the mechanism cannot be renegotiated with a positive probability on the equilibrium path.
(ii) No agent can improve upon his interim equilibrium payoff by a unilateral deviation from σ followed by renegotiation of the resulting outcome with a positive probability off the equilibrium path.

Here is an incomplete information analog of Definition 2 (Section 3.1).

Definition 6. A social choice function is *ex post renegotiation-proof* (*EPRP*) *implementable* if there exists a mechanism $\langle S, m\rangle$ that has an EPRP equilibrium σ such that the equilibrium outcome is payoff equivalent to f from the agents' interim perspective.

4.2. *Independent private values*

In this section we study environments with private values. Note that each agent's preferences over any pair of outcomes, and in particular any original outcome produced by the mechanism and any alternative outcome, are independent of his beliefs over the other agents' types and depend only on his own type.

As in the case of complete information, in environments with private values there can be no ex post inefficient equilibria that are EPRP. In every state of the world each agent knows his own preferences, and thus everyone will consent to a Pareto improving change.[35] This implies the following lemma.

---

[34] A closely related notion is that of "posterior efficiency", which is due to Forges [17,18]. An equilibrium of a given mechanism is posterior efficient if the outcomes obtained from equilibrium play cannot be renegotiated with a positive probability on the equilibrium path, where the approach to modeling renegotiation of an outcome is very similar to ours. Hence the definition of a posterior efficient equilibrium is similar to part (i) of our Definition 5. However, the concept of posterior efficiency assumes that ex post renegotiation is not anticipated by the agents when they play in the mechanism and thus cannot undermine the incentive compatibility of the original equilibrium. Thus the definition of posterior efficient equilibrium does not have an analog to part (ii) of our Definition 5.

[35] There is a similar result for posterior efficient equilibria in Forges [17,18].

Lemma 2. *In an incomplete information environment with private values every EPRP equilibrium is ex post efficient.*

For the rest of this section we restrict attention to the case where the agents' private information is distributed independently. The case of correlated private information is discussed in Section 4.3.

Any mechanism that is ex post efficient under incomplete information has to induce the agents to reveal their types. For environments with independent private values the following class of mechanisms received special attention in the literature. Each agent i reports his own type $\theta_i \in \Theta_i$, an ex post efficient alternative given the reports is chosen, and each agent is given a transfer (as a function of the reports) that equals the expected externality that the agent imposes on the other agents. Such mechanisms may be called *Groves in expectation mechanisms*, because the interim expected equilibrium transfers of each agent under such a mechanism is identical to those of some Groves mechanism.[36] Each such mechanism has an equilibrium where the agents tell the truth, and thus Groves in expectation mechanisms can be used to implement the following class of social choice functions.

Definition 7. A social choice function f is called *Groves in expectation* if $f(\theta) = (a^*(\theta), t(\theta))$ where a^* is ex post efficient and transfers t satisfy

$$E_{\theta_{-i}}\big[t_i(\theta_i, \theta_{-i})\big] = E_{\theta_{-i}}\Bigg[\sum_{j\neq i} v_j\big(a^*(\theta_i, \theta_{-i}), \theta_j\big)\Bigg] + H_i \quad \forall i \in N,\ \theta_i \in \Theta_i, \tag{4}$$

for some $H_i \in \mathbb{R}$.

Next we describe the class of social choice functions that can be implemented in a renegotiation-proof way. As in the case of complete information with two agents, the requirements of ex post renegotiation-proofness restrict the relative size of the transfers across different states of the world. We show that Groves in expectations social choice functions that are budget balanced can always be implemented in a renegotiation-proof way, and that no other social choice function can be EPRP implemented by an equilibrium of a mechanism that reveals "enough information".

Proposition 3. *Consider an incomplete information environment with independent private values.*

(i) *Every budget balanced Groves in expectation social choice function is EPRP implementable.*
(ii) *Suppose that a social choice function f is EPRP implementable by an equilibrium σ in a mechanism $\langle S, m\rangle$ with the property that for every agent i with nontrivial private information* ($\#|\Theta_i| > 1$)*, after $\langle S, m\rangle$ has been played, i's belief about the other agents is concentrated on a single type profile θ_{-i}. Then f is budget balanced Groves in expectations.*

The idea of the proof of part (i) of the result is similar to that of part (i) of Proposition 1 (Section 3.2). Consider a budget balanced Groves in expectations mechanism that corresponds

[36] The class of Groves in expectation mechanisms includes Groves mechanisms, AGV mechanisms (Arrow [3], d'Aspremont and Gérard-Varet [4]), as well as other mechanisms.

to a given budget balanced Groves in expectations social choice function. There is a truthful equilibrium of this mechanism that is ex post efficient, and that has the property that no agent's report imposes "harmful direct externalities" on the other agents (in expectation). Specifically, given the information of agent i, the expected sum of the other agents' payoffs (without renegotiation) is independent of i's report.[37] Hence, an agent who misreports his type and consequently renegotiates the resulting outcomes cannot get more than his equilibrium payoff, because the other agents must be paid at least their equilibrium payoffs, and the total surplus is maximized in equilibrium.[38]

Note that the property mentioned in part (ii) of the result is trivially satisfied in environments where only one agent has private information. Furthermore, in such environments budget balanced Groves in expectations social choice functions are simply budget balanced Groves. Hence, part (ii) of the result implies that in environments with one sided private information every EPRP implementable social choice function must be budget balanced Groves. To see the intuition, note that by Lemma 2 only ex post efficient social choice functions can be potentially EPRP implementable, and that to achieve efficiency one must rely on a report of the agent with private information. Then, by the same logic as in Example 1 (Section 3.2), renegotiation-proofness requires the reporting agent to be the residual claimant of the total surplus, or, equivalently, that the sum of the payoffs of the other agents be independent of his report. Otherwise he could report the type that minimizes the other agents' status quo payoffs, and renegotiate the outcome to an ex post efficient one while keeping all renegotiation surplus to himself.

Next, suppose that several agents have private information, and that after the mechanism is played each agent with private information can figure out the state of the world. Then, similarly to the case of one sided private information, renegotiation-proofness requires that each reporting agent fully internalizes the effect of his report on the social outcome. However, since the agents submit their messages before observing the types of the other agents, it is enough to assign them Groves in expectation transfers that would make them residual claimants in expectation, conditional on their private information at the interim stage.

Finally, suppose that several agents have private information, and that some residual uncertainty about the state of the world remains after the mechanism is played. The presence of asymmetric information at the renegotiation stage may reduce the profitability of misreporting followed by renegotiation, because even in the most favorable renegotiation scenario a deviating agent may fail to capture the entire difference between the efficient total surplus and the status quo payoffs of the other agents. In such cases the renegotiation-proofness constraints are weaker, and it may be possible that social choice functions that are different from budget balanced Groves in expectation functions are EPRP implementable.

For illustration we consider two buyer–seller examples with two sided private information. First we revisit the setting of Example 2 (Section 3.2) under incomplete information.

---

[37] Suppose agent i reports the type θ_i, and the other agents report their types truthfully. Then the expected sum of the other agents' payoffs is

$$E_{\theta_{-i}}\left[\sum_{j\neq i} v_j\big(a^*(\theta_i,\theta_{-i}),\theta_j\big)+\sum_{j\neq i} t_j(\theta_i,\theta_{-i})\right]=E_{\theta_{-i}}\left[\sum_{j\neq i} v_j\big(a^*(\theta_i,\theta_{-i}),\theta_j\big)-t_i(\theta_i,\theta_{-i})\right]=-H_i$$

where the first equality is due to budget balance, and the second follows from (4).

[38] Ausubel and Cramton [5] obtain a result similar to part (i) of Proposition 3 for multi-unit auction environments. They show that the truthful bidding remains an equilibrium of the Vickrey auction even in the presence of multiple resale opportunities.

Example 3. A buyer and a seller can trade a single good. The buyer values the good at V which can be either 0 or 2 with equal probability, and the seller values the good at C which can be either 1 or 3 with equal probability, and V and C are independently distributed. Unlike in Example 2 each agent privately observes the realization of his value. Every budget balanced Groves in expectation social choice function prescribes trade if and only if $(V, C) = (2, 1)$, the buyer's transfers satisfy $t_b(2, 1) - t_b(0, 3) = -\frac{3}{2}$ and $t_b(2, 3) - t_b(0, 1) = \frac{1}{2}$, and the seller's transfers can be found from the budget balance conditions, namely: $t_s(V, C) = -t_b(V, C)$ for every (V, C). By part (i) of Proposition 3 such social choice functions are EPRP implementable. It is possible to show that in this environment there are no other EPRP implementable social choice functions.[39]

The next example exhibits EPRP implementable social choice functions that are not budget balanced Groves in expectations.

Example 4. Consider the same setting as in Example 3, but suppose V is either 0 or 3, and C is either 1 or 2. Then every budget balanced Groves in expectation social choice function prescribes trade if and only if $V = 3$, the buyer's transfers satisfy $t_b(3, 1) - t_b(0, 2) = t_b(3, 2) - t_b(0, 1) = -\frac{3}{2}$, and the seller's transfers are the opposite of the buyer's transfers. By part (i) of Proposition 3 such social choice functions are EPRP implementable, but, as we show below, there are others as well.

Suppose the buyer is asked to report his value: after a report "$V = 3$" the good is transferred from the seller to the buyer at a price $p \in (1, 2)$, and after a report "$V = 0$" there is no trade and no payment. The buyer has a dominant strategy to tell the truth, and there will be no renegotiation on the equilibrium path because the outcome is ex post efficient. Suppose the buyer with $V = 3$ reports "$V = 0$" which results in no trade and no payment. If the seller's cost was known, then the most profitable renegotiation scenario for the buyer would be to renegotiate to trade and pay the seller no more than his cost, which would result in an expected payoff of $\frac{1}{2}(3-1) + \frac{1}{2}(3-2) = \frac{3}{2}$. However, the seller's cost is his private information, and it is not revealed in the course of the play of the mechanism. Hence, to induce both types of the seller to trade, the buyer has to pay at least 2, which results in an expected payoff of at most $3 - 2 = 1$, and this is less than the equilibrium payoff $3 - p$. To induce only the seller with $C = 1$ to trade, the buyer has to pay at least 1, which results in an expected payoff of at most $\frac{1}{2}(3 - 1) = 1 < 3 - p$. The argument for the buyer with $V = 0$ is similar.

4.3. Correlated private values

In mechanism design theory the case of incomplete information with correlated agents' beliefs is very similar to the complete information case. In the latter case the mechanism designer can ask the agents to report the state of the world and verify their reports against each other. In the former case the mechanism designer can ask the agents to report their types and "stochastically

[39] If such a social choice function existed, then by part (ii) of Proposition 3 an equilibrium of a mechanism that implements it must preserve some asymmetry of information between the agents after the play of the mechanism. However, it can be shown that even if the other agents' private information is not automatically revealed, it may nonetheless be possible for a deviating agent to extract the other agents' private information at no cost, renegotiate the trade decision to an efficient one, and capture the whole renegotiation surplus. For details see Lemma 3 in the working paper version of this paper.

compare" them against each other. For our next result we introduce the following condition from Kosenok and Severinov [26].

Condition 3. *The prior probability distribution P over Θ satisfies*:

(i) *For every agent i and every type θ_i'*:

$$P\left(\cdot \mid \theta_i'\right) \neq \sum_{\theta_i \in \Theta_i} c_{\theta_i,\theta_i'} P(\cdot \mid \theta_i) \quad \forall c_{\theta_i,\theta_i'} \in \mathbb{R}_+,\ \theta_i, \theta_i' \in \Theta_i.$$

(ii) *For every probability distribution $P' \neq P$ there exists agent i of type θ_i' (with a positive P'-probability) such that*:

$$P'\left(\cdot \mid \theta_i'\right) \neq \sum_{\theta_i \in \Theta_i} c_{\theta_i,\theta_i'} P(\cdot \mid \theta_i) \quad \forall c_{\theta_i,\theta_i'} \in \mathbb{R}_+,\ \theta_i, \theta_i' \in \Theta_i.$$

Part (i) of Condition 3 requires the conditional beliefs of each type of each agent to be linearly independent (in a restricted sense). It allows to design personalized transfers for each agent (as a function of reports) that ensure truthtelling.[40] If we want the system of personalized transfers for the agents to satisfy budget balance, then this may distort the agents' incentives to tell the truth. Part (ii) of Condition 3 takes care of this issue.

This condition allows us to obtain the following incomplete information analog of the "if" part of Proposition 2.

Proposition 4. *Consider a correlated private values environment with $n \geqslant 3$ agents with a prior distribution P that satisfies Condition 3. A social choice function f is EPRP implementable if $f(\theta) = (a(\theta), t(\theta))$ is ex post efficient for every θ.*

4.4. Discussion

Comparison with the standard approach. As in the case of complete information, ex post renegotiation under incomplete information is usually assumed to proceed according to some process that is commonly known among the agents, and is also known to the mechanism designer. For example, Beaudry and Poitevin [7] study a model with two agents who first play a game induced by a given mechanism, and then have a chance to renegotiate the realized outcomes as follows. One of the agents proposes a new mechanism; if the other agent agrees, then they play the new mechanism which determines the final outcome; otherwise the outcome prescribed by the original mechanism is implemented. The proposer of the new mechanism takes into account the prescribed outcome, his beliefs about the state of the world, beliefs about the opponent's beliefs, etc. The equilibrium of the subsequent game (consisting of the ratification procedure and the new mechanism) determines new outcomes and revised beliefs for every state of the world. This procedure uniquely determines the renegotiation mapping from the outcomes, states and the agents' beliefs into outcomes and the agents' beliefs.

In contrast we assume that there are multiple possible renegotiation possibilities, and that the mechanism designer is uncertain about which one the agents will follow. Such a mechanism

[40] See, for example, Crémer and McLean [13].

designer would presumably be interested in a mechanism that implements the desired social choice function that is robust against all reasonable renegotiation possibilities, if it exists.

Existence of EPRP mechanisms and social choice functions. In incomplete information environments with independent private values Groves in expectations mechanisms implement ex post efficient alternatives in Bayesian Nash equilibrium, and are the only such mechanisms under condition that each agent's set of types is a connected open subset of an Euclidean space and payoffs are continuously differentiable in types.[41] On the other hand, Proposition 3 shows that EPRP implementable social choice functions contain, and may be equivalent to, budget balanced Groves in expectations social choice functions, whether or not the aforementioned technical condition is satisfied. Thus, EPRP implementable social choice functions exist, since Groves in expectations mechanisms are known to exist.[42] It should be noted, however, that budget balanced Groves in expectations social choice functions do not always satisfy interim individual rationality.[43] Whenever this is the case, the existence of interim individually rational EPRP implementable social choice functions and mechanisms is not guaranteed.[44]

In the case of correlated beliefs considered in Section 4.3 EPRP implementable social choice functions were shown to exist. However, we do not know whether existence can be ensured in general.

Interdependent values. We have two points to make for this case. First, in environments with interdependent valuations ex post renegotiation-proofness no longer implies ex post efficiency. Suppose that every agent prefers the outcome (a', t') to the outcome (a, t) when the state of the world is $\theta = (\theta_1, \ldots, \theta_n)$. Nevertheless, type θ_i of agent i may be reluctant to switch to (a', t') if there exists another state $(\theta_i, \tilde{\theta}_{-i})$ in which he prefers (a, t) to (a', t'). The next example illustrates this possibility.

Example 5. A buyer and a seller can trade a single good. There are two equally likely states: in state L the buyer's value is 0 and the seller's value is 1, and in state H the buyer's value is 2 and the seller's value is $\frac{3}{2}$. The seller privately observes the realization of the state, while the buyer is uninformed. Consider a mechanism that prescribes no trade and no payment (independently of the agents' messages). The trivial equilibrium of this mechanism is EPRP despite being ex post inefficient in state H.

Suppose the agents are offered to renegotiate this outcome to trade at a price $\hat{p}$. Note that $\hat{p}$ must be at least $\frac{3}{2}$ to give the seller incentives to trade in state H. But then the seller would like to trade in state L as well. So, no trade will take place since the buyer is not willing to pay more than his expected value, which is equal to $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 2 = 1$.

---

41 See, for example, Williams [41].

42 This follows from the existence of AGV mechanisms mentioned in footnote 36.

43 This has been shown, for example, in the case of bilateral trade (Myerson and Satterthwaite [32]), public good provision (Mailath and Postlewaite [28]), and litigation and settlement (Spier [38], Klement and Neeman [25]).

44 In Example 3 from Section 4.2 no budget balanced Groves in expectation social choice function (and thus no EPRP implementable social choice function) satisfies interim individual rationality. To see this note that the interim expected payoffs of the buyer with $V = 0$ and of the seller with $C = 3$ are $\frac{1}{2}t_b(0, 1) + \frac{1}{2}t_b(0, 3)$ and $\frac{1}{2}t_s(0, 3) + \frac{1}{2}t_s(2, 3)$, respectively. Summing up these payoffs and using budget balance and the restrictions on the transfers, it follows that $\frac{1}{2}t_b(0, 1) - \frac{1}{2}t_b(2, 3) = -\frac{1}{4}$. Hence, the payoffs of the buyer with $V = 0$ and of the seller with $C = 3$ cannot both be nonnegative.

Second, if values are interdependent and correlated in the sense that they satisfy Condition 3 above, then Proposition 4 continues to hold. The argument is essentially identical to the one given for the private values case.

What next? If the EPRP requirements appear to be too stringent for a particular problem, one could use alternative approaches 1 through 4 from Section 3.3.

Another approach that could be useful is to restrict the allowed renegotiation mappings to be "credible", so that the resulting post-renegotiation state contingent allocations could not be renegotiated again in some states of the world.[45] While it is possible to show that a restriction to credible renegotiation proposals does not alter our results in the case of complete information, this is not the case under incomplete information. We have chosen not to restrict attention to "credible" renegotiation proposals, because we think that a more permissive notion of renegotiation goes well with our goal of determining which equilibria of which mechanisms are robust against various renegotiation scenarios. For example, in some cases the agents might have time for just one round of renegotiation, and then some "non-credible" renegotiation agreements may be realized.

Throughout the paper we have assumed that after a mechanism is played the realized outcome is revealed to the players. In principle there may be other possible disclosure policies. Some mechanisms may in addition reveal some or all of the actions taken by the players in the course of the play of the mechanism. Other mechanisms may reveal to each player only part of the information about the realized outcome, say, only his own transfer and some minimal information about the chosen social alternative. We conjecture that Propositions 3 and 4 would continue to hold true under any disclosure policy.

Appendix A

Proof of Proposition 1. (i) Suppose $f(\theta) = (a(\theta), t(\theta))$ is ex post efficient in every $\theta \in \Theta$, i.e. $a(\theta) = a^*(\theta)$ and $t(\theta)$ is budget balanced, and transfers t satisfy (1). Consider a mechanism where each agent i reports his own payoff relevant type $\theta_i \in \Theta_i$, and the outcome rule as a function of reports is given by f. In this mechanism truthtelling is a dominant strategy, since for each agent i of type θ_i

$$\begin{aligned} v_i\big(a^*(\theta_i,\theta_j),\theta_i\big) + t_i(\theta_i,\theta_j) &= v_i\big(a^*(\theta_i,\theta_j),\theta_i\big) + v_j\big(a^*(\theta_i,\theta_j),\theta_j\big) + H_i(\theta_j) \\ &\geqslant v_i\big(a^*\big(\theta_i',\theta_j\big),\theta_i\big) + v_j\big(a^*\big(\theta_i',\theta_j\big),\theta_j\big) + H_i(\theta_j) \\ &= v_i\big(a^*\big(\theta_i',\theta_j\big),\theta_i\big) + t_i\big(\theta_i',\theta_j\big) \quad \forall \theta_i' \neq \theta_i \end{aligned}$$

where the equalities use (1), and the inequality uses the fact that $a^*(\theta_i,\theta_j)$ is ex post efficient in state (θ_i,θ_j). Hence, the truthful equilibrium of this mechanism implements f.

[45] Such an approach is considered, for example, in Krasa [27] who studied a different notion of ex post renegotiation-proofness called *unimprovability*. Unimprovability is a property of a state contingent allocation rather than a property of an equilibrium of some mechanism, and the process by which the agents arrive at a given allocation is not modeled. The allocation is called unimprovable if the agents do not want to change it by re-trading goods or revealing further information. In this respect the definition of unimprovable allocation is related to the requirement on EPRP equilibrium outcomes given in part (i) of our Definition 5.

Suppose the above equilibrium is not EPRP. First, note that there will be no renegotiation on the equilibrium path since the outcome $(a^*(\theta), t(\theta))$ is ex post efficient for every θ. Suppose agent 1 in state $\theta = (\theta_1, \theta_2)$ can improve upon his truthtelling payoff by sending a wrong report $\theta_1' \neq \theta_1$ and consequent renegotiation of the outcome from $(a^*(\theta_1', \theta_2), t(\theta_1', \theta_2))$ to another outcome $(\hat{a}, \hat{t})$. This strategy is beneficial for agent 1 if

$$v_1(\hat{a}, \theta_1) + \hat{t}_1 > v_1\big(a^*(\theta_1, \theta_2), \theta_1\big) + t_1(\theta_1, \theta_2), \tag{5}$$

agent 2 agrees to such renegotiation if

$$v_2(\hat{a}, \theta_2) + \hat{t}_2 \geqslant v_2\big(a^*\big(\theta_1', \theta_2\big), \theta_2\big) + t_2\big(\theta_1', \theta_2\big). \tag{6}$$

Sum up (5) and (6):

$$\begin{aligned} &v_1(\hat{a}, \theta_1) + v_2(\hat{a}, \theta_2) + \hat{t}_1 + \hat{t}_2 \\ &\quad > v_1\big(a^*(\theta_1, \theta_2), \theta_1\big) + v_2\big(a^*\big(\theta_1', \theta_2\big), \theta_2\big) + t_1(\theta_1, \theta_2) + t_2\big(\theta_1', \theta_2\big) \\ &\quad = v_1\big(a^*(\theta_1, \theta_2), \theta_1\big) + v_2\big(a^*\big(\theta_1', \theta_2\big), \theta_2\big) + t_1(\theta_1, \theta_2) - t_1\big(\theta_1', \theta_2\big) \\ &\quad = v_1\big(a^*(\theta_1, \theta_2), \theta_1\big) + v_2\big(a^*(\theta_1, \theta_2), \theta_2\big) \end{aligned}$$

where the first equality uses budget balance ($t_2(\theta_1', \theta_2) = -t_1(\theta_1', \theta_2)$), and the second equality uses (1). But $a^*(\theta_1, \theta_2)$ is ex post efficient in state (θ_1, θ_2) and $\hat{t}_1 + \hat{t}_2 \leqslant 0$:

$$v_1\big(a^*(\theta_1, \theta_2), \theta_1\big) + v_2\big(a^*(\theta_1, \theta_2), \theta_2\big) \geqslant v_1(\hat{a}, \theta_1) + v_2(\hat{a}, \theta_2) + \hat{t}_1 + \hat{t}_2$$

which gives a contradiction. Hence, the truthful equilibrium of this mechanism is EPRP.

(ii) Suppose a mechanism $\langle S, m\rangle$ has an EPRP equilibrium σ that implements the social choice function f. For every state $\theta = (\theta_1, \theta_2) \in \Theta$ define a maximized total surplus: $S(\theta_1, \theta_2) = \sum_{i=1}^{2} v_i(a^*(\theta_1, \theta_2), \theta_i)$.

First, by Lemma 1 note that $f(\theta) = (a(\theta), t(\theta))$ must be ex post efficient: $a(\theta) = a^*(\theta)$ and $t_1(\theta) + t_2(\theta) = 0$ for every θ. Next, suppose agent 1 in state θ uses strategy $\sigma_1(\theta')$ where $\theta' = (\theta_1', \theta_2') \in \Theta$. Since σ is EPRP, agent 1 cannot improve upon his equilibrium payoff even if the resulting outcomes are renegotiated in the most profitable way. That is, every outcome $(\tilde{a}, \tilde{t})$ is renegotiated to an ex post efficient outcome $(a^*(\theta), \hat{t})$ such that agent 2 receives a payoff that leaves him indifferent between the two outcomes:

$$v_2\big(a^*(\theta), \theta_2\big) + \hat{t}_2 = v_2(\tilde{a}, \theta_2) + \tilde{t}_2.$$

The payoff of agent 1 in this case is

$$v_1\big(a^*(\theta), \theta_1\big) - \hat{t}_2 = v_1\big(a^*(\theta), \theta_1\big) + v_2\big(a^*(\theta), \theta_2\big) - \big(v_2(\tilde{a}, \theta_2) + \tilde{t}_2\big).$$

The renegotiation-proofness constraint to prevent such deviations by agent 1 is

$$\begin{aligned} v_1\big(a^*(\theta), \theta_1\big) + t_1(\theta) \geqslant\ & v_1\big(a^*(\theta), \theta_1\big) + v_2\big(a^*(\theta), \theta_2\big) \\ & - E_{m\circ(\sigma_1(\theta'), \sigma_2(\theta))}\big[v_2(\tilde{a}, \theta_2) + \tilde{t}_2\big] \end{aligned} \tag{7}$$

where the expectation for the last term is taken with respect to a probability measure over outcomes $(\tilde{a}, \tilde{t})$ induced by the outcome rule m and the strategy profile $(\sigma_1(\theta'), \sigma_2(\theta))$. Rewrite (7):

$$t_1(\theta) + E_{m\circ(\sigma_1(\theta'), \sigma_2(\theta))}[\tilde{t}_2] \geqslant v_2\big(a^*(\theta), \theta_2\big) - E_{m\circ(\sigma_1(\theta'), \sigma_2(\theta))}\big[v_2(\tilde{a}, \theta_2)\big] \quad \forall \theta' \neq \theta. \tag{8}$$

Repeat the same argument for agent 2:

$$t_2(\theta) + E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}[\tilde{t}_1] \geqslant v_1\big(a^*(\theta), \theta_1\big) - E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}\big[v_1(\tilde{a}, \theta_1)\big] \quad \forall \theta' \neq \theta. \tag{9}$$

Sum up (8) and (9), and use $t_1(\theta) + t_2(\theta) = 0$:

$$\begin{aligned} &E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}[\tilde{t}_1] + E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}[\tilde{t}_2] \\ &\quad \geqslant S(\theta_1, \theta_2) - E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}\big[v_1(\tilde{a}, \theta_1)\big] - E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}\big[v_2(\tilde{a}, \theta_2)\big]. \end{aligned} \tag{10}$$

Switch the roles of θ and θ' in (10):

$$\begin{aligned} &E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}[\tilde{t}_1] + E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}[\tilde{t}_2] \\ &\quad \geqslant S\big(\theta_1', \theta_2'\big) - E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}\big[v_1\big(\tilde{a}, \theta_1'\big)\big] - E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}\big[v_2\big(\tilde{a}, \theta_2'\big)\big]. \end{aligned} \tag{11}$$

Sum up (10) and (11):

$$\begin{aligned} &E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}[\tilde{t}_1 + \tilde{t}_2] + E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}[\tilde{t}_1 + \tilde{t}_2] \\ &\quad \geqslant S(\theta_1, \theta_2) + S\big(\theta_1', \theta_2'\big) - E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}\big[v_1(\tilde{a}, \theta_1) + v_2\big(\tilde{a}, \theta_2'\big)\big] \\ &\qquad - E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}\big[v_1\big(\tilde{a}, \theta_1'\big) + v_2(\tilde{a}, \theta_2)\big]. \end{aligned} \tag{12}$$

Note that the left-hand side of (12) is nonpositive since transfers $\tilde{t}$ are feasible. Using this fact and rearranging we get

$$\begin{aligned} &E_{m\circ(\sigma_1(\theta),\sigma_2(\theta'))}\big[v_1(\tilde{a}, \theta_1) + v_2\big(\tilde{a}, \theta_2'\big)\big] + E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}\big[v_1\big(\tilde{a}, \theta_1'\big) + v_2(\tilde{a}, \theta_2)\big] \\ &\quad \geqslant S(\theta_1, \theta_2) + S\big(\theta_1', \theta_2'\big). \end{aligned} \tag{13}$$

Thus a necessary condition for ex post renegotiation-proofness is

$$S\big(\theta_1, \theta_2'\big) + S\big(\theta_1', \theta_2\big) \geqslant S(\theta_1, \theta_2) + S\big(\theta_1', \theta_2'\big). \tag{14}$$

A reverse inequality in (14) is obtained if we repeat the same argument for states (θ_1, θ_2') and (θ_1', θ_2) instead of (θ_1, θ_2) and (θ_1', θ_2'). Hence, a necessary condition for ex post renegotiation-proofness is

$$S\big(\theta_1, \theta_2'\big) + S\big(\theta_1', \theta_2\big) = S\big(\theta_1', \theta_2'\big) + S(\theta_1, \theta_2) \tag{15}$$

and thus all inequalities used to arrive at (15) must hold as equalities. In particular, the left side of (13) must be equal to the left side of (14), which implies that $m \circ (\sigma_1(\theta), \sigma_2(\theta'))$ is equal to $a^*(\theta_1, \theta_2')$ with probability one, and $m \circ (\sigma_1(\theta'), \sigma_2(\theta))$ is equal to $a^*(\theta_1', \theta_2)$ with probability one. Thus (8) becomes

$$t_1(\theta) + E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}[\tilde{t}_2] = v_2\big(a^*(\theta), \theta_2\big) - v_2\big(a^*\big(\theta_1', \theta_2\big), \theta_2\big) \quad \forall \theta' \neq \theta, \tag{16}$$

and if we switch the roles of θ and θ' in (9):

$$t_2\big(\theta'\big) + E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}[\tilde{t}_1] = v_1\big(a^*\big(\theta'\big), \theta_1'\big) - v_1\big(a^*\big(\theta_1', \theta_2\big), \theta_1'\big) \quad \forall \theta \neq \theta'. \tag{17}$$

Take $\theta = (\theta_1, \theta_2)$ and $\theta' = (\theta_1', \theta_2)$, and add up (16) and (17):

$$\begin{aligned} &t_1(\theta_1, \theta_2) + t_2\big(\theta_1', \theta_2\big) + E_{m\circ(\sigma_1(\theta'),\sigma_2(\theta))}[\tilde{t}_1 + \tilde{t}_2] \\ &\quad = v_2\big(a^*(\theta_1, \theta_2), \theta_2\big) - v_2\big(a^*\big(\theta_1', \theta_2\big), \theta_2\big). \end{aligned} \tag{18}$$

Another implication of the argument above is that all inequalities used to arrive at (15) must hold as equalities is $\tilde{t}_1 + \tilde{t}_2 = 0$. Rewrite (18) using budget balance ($t_2(\theta_1', \theta_2) = -t_1(\theta_1', \theta_2)$):

$$t_1(\theta_1, \theta_2) - t_1(\theta_1', \theta_2) = v_2(a^*(\theta_1, \theta_2), \theta_2) - v_2(a^*(\theta_1', \theta_2), \theta_2).$$

Hence, the transfer rule satisfies (1). The argument for agent 2 is identical. □

Proof of Lemma 2. Let σ be an EPRP equilibrium of a mechanism $\langle S, m\rangle$. Suppose σ is not ex post efficient in some state $\theta = (\theta_1, \dots, \theta_n)$. That is, there exists some outcome (a, t) reached with positive probability by σ when the profile of types is θ, such that (a, t) is not ex post efficient in state θ.

Suppose the following indirect renegotiation mechanism is offered to the agents after outcome (a, t) is reached. The agents simultaneously announce Y or N. If at least one agent has announced N, then outcome (a, t) is realized. If all agents have announced Y, then outcome $(a^*(\theta), \hat{t})$ is realized, where $a^*(\theta)$ maximizes the total surplus in state θ, and transfers $\hat{t}$ are defined as

$$\hat{t}_i = v_i(a, \theta_i) - v_i(a^*(\theta), \theta_i) + t_i + \frac{1}{n}\left(\sum_{j=1}^{n} v_j(a^*(\theta), \theta_j) - \sum_{j=1}^{n} v_j(a, \theta_j)\right).$$

Transfers $\hat{t}$ are feasible since transfers t are feasible: $\sum_{i=1}^{n} \hat{t}_i = \sum_{i=1}^{n} t_i \leqslant 0$.

Note that agent i of type θ_i strictly prefers $(a^*(\theta), \hat{t})$ to (a, t) for any beliefs about the opponents' types:

$$\begin{aligned} v_i(a^*(\theta), \theta_i) + \hat{t}_i &= v_i(a, \theta_i) + t_i + \frac{1}{n}\left(\sum_{j=1}^{n} v_j(a^*(\theta), \theta_j) - \sum_{j=1}^{n} v_j(a, \theta_j)\right) \\ &> v_i(a, \theta_i) + t_i. \end{aligned}$$

Hence, there exists an equilibrium of this renegotiation mechanism such that in state θ all agents choose Y, and thus the outcome is $(a^*(\theta), \hat{t})$. Using a standard argument it is straightforward to construct an outcome equivalent incentive compatible direct mechanism ρ. By construction, this mechanism ρ is posterior individually rational with respect to (a, t), and it satisfies the requirements of part (i) of Definition 3.

This gives a contradiction, and hence an EPRP equilibrium σ must be ex post efficient. □

Proof of Proposition 3. (i) Suppose $f(\theta) = (a(\theta), t(\theta))$ is ex post efficient in every $\theta \in \Theta$, i.e. $a(\theta) = a^*(\theta)$ and $t(\theta)$ is budget balanced, and transfers t satisfy (4). Consider a direct revelation mechanism where each agent i reports his type $\theta_i \in \Theta_i$, and the outcome rule as a function of reports is given by f. There is a truthtelling equilibrium of this mechanism, since for each agent i of type θ_i

$$\begin{aligned} &E_{\theta_{-i}}\left[v_i(a^*(\theta_i, \theta_{-i}), \theta_i) + t_i(\theta_i, \theta_{-i})\right] \\ &\quad = E_{\theta_{-i}}\left[v_i(a^*(\theta_i, \theta_{-i}), \theta_i) + \sum_{j \neq i} v_j(a^*(\theta_i, \theta_{-i}), \theta_j)\right] + H_i \\ &\quad \geqslant E_{\theta_{-i}}\left[v_i(a^*(\theta_i', \theta_{-i}), \theta_i) + \sum_{j \neq i} v_j(a^*(\theta_i', \theta_{-i}), \theta_j)\right] + H_i \\ &\quad = E_{\theta_{-i}}\left[v_i(a^*(\theta_i', \theta_{-i}), \theta_i) + t_i(\theta_i', \theta_{-i})\right] \quad \forall \theta_i' \neq \theta_i \end{aligned}$$

where the equalities use (4), and the inequality uses the fact that $a^*(\theta_i, \theta_{-i})$ is ex post efficient in state (θ_i, θ_{-i}). Hence, the truthful equilibrium of this mechanism implements f.

Suppose the truthtelling equilibrium σ is not EPRP. First, suppose that some outcome (a, t) is reached with positive probability by the equilibrium play, and it can be renegotiated with positive probability. Note that the beliefs of each agent i over Θ_{-i} are derived by Bayes' rule from P conditional on (a, t), σ, and θ_i. Thus we can say that at this point the agents share a common "prior" over Θ, which is derived by Bayes' rule from P conditional on (a, t) and σ. Also note that it is commonly known among the agents that (a, t) is ex post efficient. Hence, by the "no-trade theorem" (Milgrom and Stokey [30]) there does not exist a feasible incentive compatible mapping ρ that improves upon outcome (a, t) as required by part (i) of Definition 3.

Next, suppose agent 1 of type θ_1 can improve upon his interim payoff from truthtelling by sending a wrong report $\theta_1' \neq \theta_1$ and consequent renegotiation of the resulting outcomes. For example, if the true state is (θ_1, θ_{-1}), then outcome $(a^*(\theta_1', \theta_{-1}), t(\theta_1', \theta_{-1}))$ will be reached by the play of the mechanism. By part (ii) of Definition 3 every viable renegotiation mechanism $(\hat{a}, \hat{t}) : \Theta \to \Delta(A \times \mathbb{T})$ must satisfy posterior individual rationality with respect to the reached outcome for each agent $i \neq 1$

$$v_i\big(\hat{a}(\theta_1, \theta_{-1}), \theta_i\big) + \hat{t}_i(\theta_1, \theta_{-1}) \geqslant v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big) + t_i\big(\theta_1', \theta_{-1}\big). \tag{19}$$

Sum up (19) over $i \neq 1$ and take expectation over Θ_{-1}:

$$\begin{aligned} &E_{\theta_{-1}}\left[\sum_{i\neq 1} v_i\big(\hat{a}(\theta_1, \theta_{-1}), \theta_i\big) + \sum_{i\neq 1} \hat{t}_i(\theta_1, \theta_{-1})\right] \\ &\quad \geqslant E_{\theta_{-1}}\left[\sum_{i\neq 1} v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big) + \sum_{i\neq 1} t_i\big(\theta_1', \theta_{-1}\big)\right]. \end{aligned} \tag{20}$$

Let us derive an upper bound on the interim expected payoff of agent 1 from such a strategy:

$$\begin{aligned} &E_{\theta_{-1}}\big[v_1\big(\hat{a}(\theta_1, \theta_{-1}), \theta_1\big) + \hat{t}_1(\theta_1, \theta_{-1})\big] \\ &\quad \leqslant E_{\theta_{-1}}\left[v_1\big(\hat{a}(\theta_1, \theta_{-1}), \theta_1\big) - \sum_{i\neq 1} \hat{t}_i(\theta_1, \theta_{-1})\right] \\ &\quad \leqslant E_{\theta_{-1}}\left[v_1\big(\hat{a}(\theta_1, \theta_{-1}), \theta_1\big) + \sum_{i\neq 1} v_i\big(\hat{a}(\theta_1, \theta_{-1}), \theta_i\big) \right. \\ &\qquad \left. - \sum_{i\neq 1} v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big) - \sum_{i\neq 1} t_i\big(\theta_1', \theta_{-1}\big)\right] \\ &\quad = E_{\theta_{-1}}\left[\sum_{i=1}^{n} v_i\big(\hat{a}(\theta_1, \theta_{-1}), \theta_i\big)\right] + H_i + E_{\theta_{-1}}\left[\sum_{i=1}^{n} t_i\big(\theta_1', \theta_{-1}\big)\right] \end{aligned}$$

where the first inequality uses feasibility condition for transfers ($\hat{t}_1(\theta_1, \theta_{-1}) \leqslant -\sum_{i\neq 1} \hat{t}_i(\theta_1, \theta_{-1})$), the second inequality uses (20), and the equality uses definition of Groves in expectations transfers (4). Using budget balance ($\sum_{i=1}^{n} t_i(\theta_1', \theta_{-1}) = 0$) and the fact that $a^*(\theta_1, \theta_{-1})$ is ex post efficient in state (θ_1, θ_{-1}) we have

$$\begin{aligned}
&E_{\theta_{-1}}\left[v_1\left(\hat{a}(\theta_1,\theta_{-1}),\theta_1\right)+\hat{t}_1(\theta_1,\theta_{-1})\right] \\
&\quad\leqslant E_{\theta_{-1}}\left[\sum_{i=1}^{n} v_i\left(a^*(\theta_1,\theta_{-1}),\theta_i\right)\right]+H_i \\
&\quad= E_{\theta_{-1}}\left[v_1\left(a^*(\theta_1,\theta_{-1}),\theta_1\right)+t_1(\theta_1,\theta_{-1})\right]
\end{aligned}$$

where the equality is by definition of Groves in expectations transfers (4). This gives a contradiction. Hence, the truthful equilibrium of this mechanism is EPRP.

(ii) Suppose a mechanism $\langle S, m\rangle$ has an EPRP equilibrium σ that implements the social choice function f, and satisfies the property described in the statement of the result.

First, by Lemma 2 note that $f(\theta)=(a(\theta),t(\theta))$ must be ex post efficient: $a(\theta)=a^*(\theta)$ and $\sum_{i=1}^{n} t_i(\theta)=0$ for every θ. Next, since σ is EPRP, no agent should be able to improve upon his interim equilibrium payoff by deviating from his equilibrium strategy even if the resulting outcomes are renegotiated in the most profitable way. In order to write down these renegotiation-proofness constraints, we first need to find the highest payoff that a given agent can get from renegotiation.

Specifically, suppose the state is (θ_1,θ_{-1}), and outcome (a,t) was reached by the mechanism when every agent $i\neq 1$ has played his equilibrium strategy, and agent 1 has played strategy $\sigma_1(\theta_1')$ where $\theta_1'\neq\theta_1$.[46,47] By ex post efficiency a must be equal to $a^*(\theta_1',\theta_{-1})$ with probability one, and t must be budget balanced with probability one. We will assume that $a=a^*(\theta_1',\theta_{-1})$ and $\sum_{i=1}^{n} t_i=0$ hold everywhere, and it is straightforward to adapt the argument to allow for inefficiencies to happen on zero probability events. By part (ii) of Definition 3 every viable renegotiation mechanism must satisfy posterior individual rationality with respect to the reached outcome for each agent $i\neq 1$. Thus every outcome $(\hat{a},\hat{t})$ to which renegotiation may lead to must satisfy

$$v_i(\hat{a},\theta_i)+\hat{t}_i\geqslant v_i\left(a^*\left(\theta_1',\theta_{-1}\right),\theta_i\right)+t_i. \tag{21}$$

Sum up (21) over $i\neq 1$ and rearrange

$$-\sum_{i\neq 1}\hat{t}_i\leqslant\sum_{i\neq 1}v_i(\hat{a},\theta_i)-\sum_{i\neq 1}v_i\left(a^*\left(\theta_1',\theta_{-1}\right),\theta_i\right)-\sum_{i\neq 1}t_i. \tag{22}$$

Let us derive an upper bound on the payoff of agent 1 in this case

$$\begin{aligned}
v_1(\hat{a},\theta_i)+\hat{t}_1&\leqslant v_1(\hat{a},\theta_i)-\sum_{i\neq 1}\hat{t}_i \\
&\leqslant\sum_{i=1}^{n}v_i(\hat{a},\theta_i)-\sum_{i\neq 1}v_i\left(a^*\left(\theta_1',\theta_{-1}\right),\theta_i\right)+t_1 \\
&\leqslant\sum_{i=1}^{n}v_i\left(a^*(\theta_1,\theta_{-1}),\theta_i\right)-\sum_{i\neq 1}v_i\left(a^*\left(\theta_1',\theta_{-1}\right),\theta_i\right)+t_1
\end{aligned}$$

where the first inequality uses feasibility condition for transfers ($\hat{t}_1\leqslant-\sum_{i\neq 1}\hat{t}_i$), the second inequality uses (22) and budget balance ($t_1=-\sum_{i\neq 1}t_i$), and the last inequality is by definition of $a^*(\theta_1,\theta_{-1})$.

---

46 If $\sigma_1(\theta_1')$ is a mixed strategy, then suppose agent 1 has played a particular pure strategy in its support.

47 If $\#|\Theta_1|=1$, then pick any other agent j such that $\#|\Theta_j|>1$.

Next we show that for every θ_{-1} it is possible to achieve the above upper bound on the renegotiation payoff of agent 1 of type θ_1 who has mimicked type θ_1'. The key condition which allows to show this is the assumption that the knowledge of the outcome is sufficient to compute θ_{-1} (assuming that every agent $i \neq 1$ has played his equilibrium strategy).

Specifically, suppose outcome (a, t) was reached. If the implied θ_{-1} is such that $a^*(\theta_1, \theta_{-1}) = a^*(\theta_1', \theta_{-1})$, then renegotiation cannot improve the payoff of agent 1, and the upper bound is reached without renegotiation. Otherwise the following indirect renegotiation mechanism is offered to the agents. The agents simultaneously announce Y or N. If at least one agent has announced N, then outcome (a, t) is realized. If all agents have announced Y, then outcome $(a^*(\theta_1, \theta_{-1}), \hat{t})$ is realized, where $a^*(\theta_1, \theta_{-1})$ maximizes the total surplus in state (θ_1, θ_{-1}), and transfers $\hat{t}$ are defined as

$$\begin{aligned}
\hat{t}_i &= v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big) - v_i\big(a^*(\theta_1, \theta_{-1}), \theta_i\big) + t_i \quad \forall i \neq 1, \\
\hat{t}_1 &= \sum_{i \neq 1} v_i\big(a^*(\theta_1, \theta_{-1}), \theta_i\big) - \sum_{i \neq 1} v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big) + t_1.
\end{aligned}$$

Note that transfers $\hat{t}$ are budget balanced since transfers t are budget balanced: $\sum_{i=1}^n \hat{t}_i = \sum_{i=1}^n t_i = 0$. By construction, every agent $i \neq 1$ of type θ_i is indifferent between $(a^*(\theta_1, \theta_{-1}), \hat{t})$ and (a, t) for any beliefs about the opponents' types, while agent 1 of type θ_1 strictly prefers $(a^*(\theta_1, \theta_{-1}), \hat{t})$ to (a, t).

Hence, there exists an equilibrium of this renegotiation mechanism such that in state (θ_1, θ_{-1}) all agents choose Y, and thus the outcome is $(a^*(\theta_1, \theta_{-1}), \hat{t})$. Using a standard argument it is straightforward to construct an outcome equivalent incentive compatible direct mechanism ρ. By construction, this mechanism ρ is posterior individually rational with respect to (a, t), and it satisfies the requirements of part (ii) of Definition 3.

Take an expectation of the upper bound on the renegotiation payoff of agent 1 of type θ_1 over possible outcomes (a, t) and the opponents' type profiles θ_{-1}

$$\begin{aligned}
&E_{\theta_{-1}}\left[E_{m\circ(\sigma_1(\theta_1'), \sigma_{-1}(\theta_{-1}))}\left[\sum_{i=1}^n v_i\big(a^*(\theta_1, \theta_{-1}), \theta_i\big) - \sum_{i \neq 1} v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big) + t_1\right]\right] \\
&\quad = E_{\theta_{-1}}\left[\sum_{i=1}^n v_i\big(a^*(\theta_1, \theta_{-1}), \theta_i\big) - \sum_{i=1}^n v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big)\right] \\
&\qquad + E_{\theta_{-1}}\big[v_1\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_1'\big) + t\big(\theta_1', \theta_{-1}\big)\big]
\end{aligned}$$

where the equality follows from the fact that σ implements f, and thus the resulting outcome is interim payoff equivalent to $(a^*(\theta_1', \theta_{-1}), t(\theta_1', \theta_{-1}))$. Hence, the renegotiation-proofness constraint to prevent deviations of this sort by agent 1 of type θ_1 is

$$\begin{aligned}
&E_{\theta_{-1}}\big[v_1\big(a^*(\theta_1, \theta_{-1}), \theta_1\big) + t_1(\theta_1, \theta_{-1})\big] \\
&\quad \geqslant E_{\theta_{-1}}\left[\sum_{i=1}^n v_i\big(a^*(\theta_1, \theta_{-1}), \theta_i\big) - \sum_{i=1}^n v_i\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_i\big)\right] \\
&\qquad + E_{\theta_{-1}}\big[v_1\big(a^*\big(\theta_1', \theta_{-1}\big), \theta_1'\big) + t\big(\theta_1', \theta_{-1}\big)\big]. \qquad (23)
\end{aligned}$$

Rewrite (23):

$$\begin{aligned}&E_{\theta_{-1}}\big[t_1(\theta_1,\theta_{-1})\big]-E_{\theta_{-1}}\big[t_1\big(\theta_1',\theta_{-1}\big)\big]\\&\quad\geqslant E_{\theta_{-1}}\bigg[\sum_{i\neq 1}v_i\big(a^*(\theta_1,\theta_{-1}),\theta_i\big)\bigg]-E_{\theta_{-1}}\bigg[\sum_{i\neq 1}v_i\big(a^*\big(\theta_1',\theta_{-1}\big),\theta_i\big)\bigg].\end{aligned}\tag{24}$$

Switching the roles of θ_1 and θ_1' we find that (24) must hold with equality. Thus the transfer rule satisfies (4). The argument for other agents i such that $\#|\Theta_i| > 1$ is identical. □

Proof of Proposition 4. Suppose $f(\theta) = (a(\theta), t(\theta))$ is ex post efficient in every $\theta \in \Theta$, i.e. $a(\theta) = a^*(\theta)$ and $t(\theta)$ is budget balanced. Consider a mechanism where each agent i reports his type $\theta_i \in \Theta_i$, and the outcome rule as a function of reports $\theta = (\theta_1, \ldots, \theta_n)$ specifies $\alpha(\theta) = a^*(\theta)$ and budget balanced transfers $\tau(\theta)$ such that

$$E_{\theta_{-i}|\theta_i}\big[\tau_i(\theta_i,\theta_{-i})\big]=E_{\theta_{-i}|\theta_i}\big[t_i(\theta_i,\theta_{-i})\big]\quad \forall i,\ \theta_i \tag{25}$$

and

$$E_{\theta_{-i}|\theta_i}\big[\tau_i\big(\theta_i',\theta_{-i}\big)\big]<-L\quad \forall i,\ \theta_i'\neq\theta_i \tag{26}$$

where $L \in \mathbb{R}$ is chosen to be sufficiently large. The existence of such transfer rule τ follows from the results in Kosenok and Severinov [26].[48]

There exists a truthful equilibrium since the expected penalty L is chosen to be large enough. To see that this equilibrium is EPRP first note that there will be no renegotiation on the equilibrium path since the outcome $(\alpha(\theta), \tau(\theta))$ is ex post efficient for every θ. If agent i misreports, then his expected payment is more than L, which is chosen to be large enough to make such a deviation unprofitable, even if the deviator consequently renegotiates the outcome and captures the implied increase in the total surplus. □

References

[1] P. Aghion, M. Dewatripont, P. Rey, Renegotiation design with unverifiable information, Econometrica 62 (1994) 257–282.
[2] P. Amoros, Nash implementation and uncertain renegotiation, Games Econ. Behav. 49 (2004) 424–434.
[3] K. Arrow, The property rights doctrine and demand revelation under incomplete information, in: M. Boskin (Ed.), Economics and Human Welfare, Academic Press, New York, 1979, pp. 23–39.
[4] C. d'Aspremont, L.-A. Gérard-Varet, Incentives and incomplete information, J. Public Econ. 11 (1979) 25–45.
[5] L. Ausubel, P. Cramton, Vickrey auctions with reserve pricing, Econ. Theory 23 (2004) 493–505.
[6] S. Baliga, T. Sjöström, Contracting with third parties, Amer. Econ. J.: Microeconomics 1 (2009) 75–100.
[7] P. Beaudry, M. Poitevin, Contract renegotiation: a simple framework and implications for organization theory, Can. J. Econ. 28 (1995) 302–335.
[8] D. Bergemann, S. Morris, Robust mechanism design, Econometrica 73 (2005) 1771–1813.
[9] T. Börgers, D. Smith, Robustly ranking mechanisms, Amer. Econ. Rev. 102 (2012) 325–329.
[10] Y.K. Che, D.B. Hausch, Cooperative investment and the value of contracting, Amer. Econ. Rev. 89 (1999) 125–147.
[11] T. Chung, Incomplete contracts, specific investment, and risk sharing, Rev. Econ. Stud. 58 (1992) 1031–1042.
[12] K.S. Chung, J. Ely, Foundations of dominant strategy mechanisms, Rev. Econ. Stud. 74 (2007) 447–476.

[48] The result follows from their Corollary 1 (p. 134) to Theorem 1 (pp. 132–133) if we note that: (i) the uniqueness of the ex post efficient social alternative for every state implies that the agents' equilibrium interim transfers (given by (25)) are uniquely determined; (ii) the agents' interim transfers from misreporting (given by the left side of (26)) can be made arbitrarily large (p. 149).

[13] J. Crémer, R. McLean, Full extraction of the surplus in Bayesian and dominant strategy auctions, Econometrica 56 (1988) 1247–1257.
[14] M. Dewatripont, Renegotiation and information revelation over time: the case of optimal labor contracts, Quart. J. Econ. 104 (1989) 589–619.
[15] A. Edlin, S. Reichelstein, Holdups, standard breach remedies, and optimal investment, Amer. Econ. Rev. 86 (1996) 478–501.
[16] R. Evans, Mechanism design with renegotiation and costly messages, Econometrica 80 (2012) 2089–2104.
[17] F. Forges, Some thoughts on efficiency and information, in: K. Binmore, A. Kirman, P. Tani (Eds.), Frontiers of Game Theory, MIT Press, Cambridge, MA, 1993, pp. 133–149.
[18] F. Forges, Posterior efficiency, Games Econ. Behav. 6 (1994) 238–261.
[19] D. Fudenberg, J. Tirole, Game Theory, MIT Press, Cambridge, MA, 1991.
[20] R. Garratt, T. Tröger, Speculation in standard auctions with resale, Econometrica 74 (2006) 753–769.
[21] J.R. Green, J.-J. Laffont, Incentives in Public Decision Making, North-Holland, Amsterdam, 1979.
[22] J.R. Green, J.-J. Laffont, Non verifiability, costly renegotiation and efficiency, Ann. Econ. Statist. 36 (1994) 81–95.
[23] T. Groves, Incentives in teams, Econometrica 41 (1973) 617–631.
[24] B. Holmström, Groves' scheme on restricted domains, Econometrica 47 (1979) 1137–1144.
[25] A. Klement, Z. Neeman, Against compromise: a mechanism design approach, J. Law, Econ., Organ. 21 (2005) 285–314.
[26] G. Kosenok, S. Severinov, Individually rational, budget-balanced mechanisms and allocation of surplus, J. Econ. Theory 140 (2008) 126–161.
[27] S. Krasa, Unimprovable allocations in economies with incomplete information, J. Econ. Theory 87 (1999) 144–168.
[28] G. Mailath, A. Postlewaite, Asymmetric information bargaining problems with many agents, Rev. Econ. Stud. 57 (1990) 351–367.
[29] E. Maskin, J. Moore, Implementation and renegotiation, Rev. Econ. Stud. 66 (1999) 39–56.
[30] P. Milgrom, N. Stokey, Information, trade and common knowledge, J. Econ. Theory 26 (1982) 17–27.
[31] J. Moore, Implementation, contracts, and renegotiation in environments with complete information, in: J.-J. Laffont (Ed.), Advances in Economic Theory: Sixth World Congress, vol. 1, Cambridge University Press, Cambridge, 1992, pp. 182–282.
[32] R.B. Myerson, M.A. Satterthwaite, Efficient mechanisms for bilateral trading, J. Econ. Theory 29 (1983) 265–281.
[33] W. Rogerson, Contractual solutions to the hold-up problem, Rev. Econ. Stud. 59 (1992) 777–793.
[34] A. Rubinstein, A. Wolinsky, Renegotiation-proof implementation and time preferences, Amer. Econ. Rev. 82 (1992) 600–614.
[35] I. Segal, M. Whinston, The Mirrlees approach to mechanism design with renegotiation (with applications to hold-up and risk sharing), Econometrica 70 (2002) 1–45.
[36] I. Segal, M. Whinston, Property rights, in: R. Gibbons, J. Roberts (Eds.), Handbook of Organizational Economics, Princeton University Press, Princeton and Oxford, 2012, pp. 100–158.
[37] T. Sjöström, Undominated Nash implementation with collusion and renegotiation, Games Econ. Behav. 26 (1999) 337–352.
[38] K.E. Spier, Pretrial bargaining and the design of fee shifting rules, RAND J. Econ. 25 (1994) 197–214.
[39] M. Walker, On the nonexistence of a dominant strategy mechanism for making optimal public decisions, Econometrica 48 (1980) 1521–1540.
[40] J. Watson, Contract, mechanism design, and technological detail, Econometrica 75 (2007) 55–81.
[41] S.R. Williams, A characterization of efficient, Bayesian incentive compatible mechanisms, Econ. Theory 14 (1999) 155–180.
[42] T. Yamashita, A necessary condition for implementation in undominated strategies, with applications to robustly optimal trading mechanisms, manuscript, Toulouse School of Economics, 2012.