Posterior Renegotiation-Proofness in a Two-Person Decision Problem*

Kei Kawakami†

April 5, 2014

Abstract

When two agents with private information use a mechanism to determine an outcome, what happens when they cannot commit to their actions and to a mechanism? We study this problem by allowing agents to hold on to a proposed outcome in one mechanism while they play another mechanism and learn new information. A decision rule is *posterior renegotiation-proof* if it is posterior implementable and robust to a posterior proposal of any posterior implementable decision rule. We identify a class of economies for which such decision rules exist. We also show how the inability to commit to the mechanism constrains equilibrium: a posterior renegotiation-proof decision rule must be implemented in a mechanism with at most five actions for two agents.

Keywords: Information aggregation, Limited commitment, Posterior efficiency, Posterior implementation, Renegotiation-proofness.

*I would like to thank Georgy Artemov, Chris Edmond, and seminar participants at the 2013 North American Summer Meeting of Econometric Society for helpful comments. All errors are mine. Support from the Faculty of Business and Economics at University of Melbourne is gratefully acknowledged.

†Department of Economics, University of Melbourne, e-mail: keik@unimelb.edu.au, website: http://sites.google.com/site/econkeikawakami/

1 Introduction

When two agents with private information use a mechanism to determine an outcome, what happens if they cannot commit to (i) actions and (ii) the mechanism? In this paper, we study this problem by allowing agents to hold on to a proposed outcome in one mechanism while they play another mechanism to obtain a new proposed outcome to compare. In this environment, while commitment to the first mechanism is not assumed, the first mechanism has a natural advantage: it can induce agents to reveal information before the alternative mechanism is played. Of course, information revelation cannot be arbitrary, as we assume no commitment to actions. We propose a concept of *posterior renegotiation-proof* (henceforth PRP) decision rules, which may be seen as describing the final outcome in this environment.

To study this problem, we cannot use the Revelation Principle, which assumes both types of commitment. Green and Laffont (1987) relaxed commitment to actions by proposing *posterior implementation*. Posterior implementation is stronger than Bayesian implementation, because agents' strategies must remain optimal against one another after any realization of an equilibrium action profile. Green and Laffont interpret posterior implementable decision rules as representing a communication process without binding commitment to actions. Therefore, a posterior implementable decision rule creates "an incentive compatible information structure", where it is common knowledge that no agent has an incentive to reveal new information by reconsidering his action *available in the given mechanism.*

In this paper, we take this idea one step further by relaxing commitment to a mechanism. Suppose agents played a particular posterior equilibrium. Taking the equilibrium outcome and revealed information as given, can a third party propose a new mechanism such that all agents can prefer a posterior equilibrium outcome of the new mechanism to the first outcome? Which decision rules are robust to such *posterior renegotiation*? To capture the robustness to such counter-proposals, we propose the following solution concept: a decision rule is PRP if it is (i) posterior implementable and (ii) robust to a posterior proposal of any posterior implementable decision rule. Obviously the part (ii) is far from precise and many notions of "robustness" can be considered. We take an extreme approach in this regard. A posterior implementable decision rule is "robust" if, conditional on the revealed information, there is no posterior implementable decision rule that either reveals new information or makes everyone better off without revealing information. Therefore, a PRP decision rule creates "a renegotiation-proof information structure", where it is common knowledge that there is no mechanism which leads to a new incentive compatible information structure. A PRP decision rule leaves no possibility for future improvement as long as we only allow information revelation by a public play of mechanisms without commitment to actions.

We characterize PRP decision rules in a two-person problem of Green and Laffont (1987). We identify conditions under which PRP decision rules exist. We also show how posterior renegotiation-proofness constrains equilibrium. In particular, any PRP decision rule must be implemented in a mechanism with at most five actions for two agents.

We provide two economic motivations for our solution concept. First, the renegotiation typically affects the incentive compatibility of the original decision rule, once the renegotiation is rationally anticipated. Therefore, without commitment to a mechanism, a concept of incentive compatibility may well be vacuous unless the possibility of the renegotiation is properly addressed. This problem has been well known in the literature (Holmstrom and Myerson 1983 and Forges 1994), but has proven to be difficult to analyze. One difficulty is that the analysis becomes sensitive to details of the renegotiation process. Our approach is to develop a sufficiently strong solution concept to avoid being sensitive to details of the renegotiation process, but not too strong such that the concept can have the predictive power for the final outcome. We elaborate more on this point in the next subsection.

Second, the form of renegotiation-proofness developed in this paper has some practical relevance in the context of market competition. Consider a financial service provider who offers an intermediation service for two investors (the service is a mechanism in which investors can interact). The market is regulated such that (i) a service provider cannot force the payment for its service until two parties reach a voluntary agreement (investor protection) and (ii) a communication process must be made public (disclosure regulation). Competing financial service providers may propose a new mechanism to the investors based on the information revealed in the first service. Therefore, without commitment to a mechanism, the possibility of posterior counter-proposals affects the ability of the intermediaries to design a mechanism. Our analysis indicates that to be robust to counter-proposals, the service offered in the market will be of very limited variety and, to the contrary of the intension of the regulation, the market may neither be investor-friendly nor transparent.

To our knowledge, our work is the first attempt to relax both commitments to actions and a mechanism. A disadvantage of our approach is that the characterization of PRP decision rules depends on the characterization of posterior implementation, which is still an open question for a general environment.[1] In this paper, we remain in the environment of Green and Laffont (1987). While we present new results on posterior implementation, our main contribution is the development of the new concept of PRP decision rules.

The rest of the paper is organized as follows. The next subsection offers more discussion of our solution concept and its relation to other solution concepts. Section 2 describes the model environment. Section 3 formally defines a PRP decision rule and characterizes it. Section 4 concludes. Section 5 contains proofs.

1.1 PRP and related solution concepts

We discuss key ingredients of our solution concept and compare it with other concepts in the literature. Our solution concept combines the idea of posterior implementation (robustness of *individual* optimality to information revealed *in a mechanism*) and the idea of renegotiation-proofness (robustness of *collective* optimality to information revealed *by comparing alternatives*). These two ideas were studied separately in the literature. We argue that we should consider both together because they are typically not independent each other.

[1] See Lopomo (2000), Jehiel et al (2007), Vartiainen (2013) for recent studies of posterior implementation.

Table 1. Related solution concepts.

Info revelation	*A*. In mechanism	*B*. Comparison to alternatives
1. Individual optimality	Posterior implementation Green and Laffont (1987)	
2. Collective optimality	Posterior efficiency Forges (1994)	Durability, Holmstrom and Myerson (1983) Security, Cramton and Palfrey (1995)

Table 1 shows related solution concepts that capture *robustness of optimality to endogenous information revelation* in one way or another. Commitment to actions is associated with the first row, while commitment to a mechanism is associated with the second row. Renegotiation is associated with the column B. In the literature, different solution concepts were proposed to address a different part of **Table 1**. Our solution concept can be viewed as addressing two cells (1, A) and (2, B) at the same time, with a particular timing assumption. Thus, unlike other solution concepts, our solution concept captures robustness of both types of optimality (individual and collective) to both types of information revelation (in mechanism and by comparison to alternatives). We capture the robustness of individual optimality by posterior implementation, while we make the robustness of collective optimality sufficiently strong such that the solution concept is not sensitive to details of the renegotiation process. This approach allows us to address the interdependence between incentive compatibility and renegotiation. We explain our solution concept in more detail below.

Posterior renegotiation-proofness. Any model of renegotiation and the robustness of the final outcome against it (i.e. renegotiation-proofness) requires some assumption on the *timing* of comparison and the *criteria* for the comparison. In terms of the timing, we allow agents to play two mechanisms and observe actions before making the comparison. We use a word "posterior" because the comparison is made *after* two mechanisms were played and beliefs were updated. Note, however, that we allow only endogenously revealed information by equilibrium actions to be relevant for the comparison. This timing assumption captures a natural advantage of the first mechanism: it can reveal information before the second mechanism is played. This is where the two sources of information revelation become interdependent. More specifically, we ask the following question: Can we design the first mechanism to make sure it only reveals information that "protects" itself against alternatives? The answer depends on how we compare alternatives and what "protects" means in that context.

The criteria for the comparison has an intricate effect on the nature of solution concepts, because implementing the criteria may further reveal information. To see this, consider the final comparison stage after two mechanisms were played by two agents. Given the updated beliefs, there are four relevant cases:

Case 1. Both agents of all types prefer the first outcome to the second outcome.

Case 2. Both agents of all types prefer the second outcome to the first outcome.

Case 3. One agent of all types prefers the first outcome while the other agent of all types prefers the second outcome (agents always disagree but the disagreement pattern does not depend on state).

Case 4. There is state-dependent agreement or disagreement, i.e., agents agree (disagree) only in some state, or the agreement (disagreement) pattern changes across states.

What would be a reasonable criteria to choose one outcome in each case? We argue that the first outcome should be chosen as the final outcome in Case 1 while in Case 2 the opposite should be the case, because the chosen outcome gets unanimous support and, more importantly, this decision *does not change agents' beliefs*. In Case 3, we need to impose some assumption to choose one outcome in favor of one agent. We give agents a veto power to choose the first outcome as the final outcome. This is also reasonable, because even if the identity of vetoer was made public, this decision reveals no information. So far, the reasonable criteria for the first outcome to be renegotiation-proof should be clear: there cannot be a new mechanism which can result in Case 2.

The complication arises in Case 4, where agreement/disagreement varies across states. In this case, revealing agreement/disagreement reveals new information, which may alter agreement/disagreement. Here, details of the renegotiation process matters, especially when we require the final outcome be determined in a consistent way. This issue is closely related to the equilibrium refinement in a signaling game, and the literature offers different notions of renegotiation-proofness, each corresponding to a different version of signaling games and equilibrium refinements.

In this paper, rather than proposing yet another renegotiation process in Case 4, we propose a solution concept that is robust to details of the renegotiation process. We achieve this by requiring that, after the play of the two mechanisms, Case 4 can not arise: the first mechanism must have the property that, once it is played, there cannot be a new mechanism which can result in Case 4. The information revealed in the first mechanism "protects" itself, because it eliminates the *possibility* of further endogenous information revelation. All in all, our assumptions lead to the following notion of renegotiation-proofness: a decision rule implemented by a given mechanism is renegotiation-proof if, whenever agents play another mechanism, it always leads to either Case 1 or Case 3.

Combining the idea of posterior implementability and the renegotiation-proofness developed above, a decision rule is PRP if it is posterior implementable and in every posterior equilibrium, there is no posterior implementable decision rule that is either (i) not constant or (ii) constant and makes all agents of all types weakly better off and at least one agent of some types strictly better off. In our solution concept, while agents are assumed to be passive in the selection of mechanisms, they acquire new information by playing a proposed mechanism. Therefore, (i) requires that the first mechanism must be designed such that, once equilibrium actions reveal information, it is common knowledge that there can be no more information revelation by open communication without commitment to actions. Additionally, (ii) rules out the remaining possibility of the improvement *without* information revelation. Note that (ii) does not rule out a constant rule that makes all agents better off in *some* state but not in other state. However, if this rule is actually implementable only in states that makes all agents better off, there must be information revelation. As long as this information revelation must occur in some game without commitment to actions, the resulting decision rule must be a posterior implementable rule that is NOT constant. Thus, the possibility of improvement by a constant rule that makes all agents better off only in some state is eliminated by (i). The key to satisfy "no more information revelation" is to let agents reveal a right amount of information in the first mechanism: not too much so that agents can commit to actions, but not too little so that another mechanism cannot induce further endogenous information revelation.

Related solution concepts. To study renegotiation-proofness, we build on Green and Laffont (1987). An advantage of this approach is that posterior implementation makes it explicit that a different *indirect* mechanism creates a different information structure, with respect to which a decision rule must remain individually optimal (incentive compatibility). We apply the same idea for collective optimality (renegotiation-proofness). As a result, our solution concept naturally captures an open-ended negotiation process in which new information is endogenously revealed and the third party can propose a new mechanism to induce more information revelation. Hence, it provides an insight into what the robust final outcome should be in such a negotiation process.

Our solution concept and Forges' (1994) *posterior efficiency* share the idea that comparison of two mechanisms should be based on endogenously revealed information in the first mechanism. The first difference is that in our solution concept endogenous information is due to a public action profile, while in posterior efficiency it is due to a public outcome. Hence, commitment to actions is still assumed in posterior efficiency. Two concepts are also different in the timing of comparison. In posterior efficiency, comparison is made after the first outcome is realized, but before the second mechanism is played. This leaves the possibility that after the second mechanism is played and beliefs are updated, the posterior efficient outcome might be overturned. Our solution concept is robust to this consideration. However, there is no clear inclusion relationship between two concepts. Posterior efficiency (see $\mathcal{E}$ on p. 250, Forges 1994) assumes that for the second decision rule to overrule the first outcome, it must be *Bayesian incentive compatible* and preferred to the first outcome by all agents of *all* types. First, a posterior efficient decision rule may not be posterior implementable, and hence may not be PRP. Second, a PRP decision rule is Bayesian incentive compatible. However, after the first mechanism is played, there may be another decision rule that is Bayesian incentive compatible (with respect to information revealed by the first outcome) but not posterior implementable (with respect to information revealed by actions), and also looks like an improvement (in the sense used to define $\mathcal{E}$). Therefore, a PRP decision rule may not be posterior efficient.

Security by Cramton and Palfrey (1995) and *durability* by Holmstrom and Myerson (1983) are concerned about the interim comparison of Bayesian incentive compatible decision rules, with different criteria applied to the case where agreement/disagreement reveals information (Case 4). Their focus is information revelation by the strategic participation decision and its consequence for interim renegotiation-proofness, while our focus is information revelation by the play of mechanisms and its consequence for posterior renegotiation-proofness. Because of our timing assumption and the use of posterior implementability, our solution concept highlights the informational role of the first mechanism, while in their approach the status quo mechanism itself does not reveal any information. Our solution concept may appear stronger than theirs by requiring that Case 4 should not arise. Again, however, because different endogenous information is used to compare two mechanisms in each approach, there is no obvious inclusion relationship.

2 Model Environment

Our model is identical to Green and Laffont's model (1987) of a collective decision problem. Two agents i, j with type $\theta = (\theta_i, \theta_j) \in \Theta_i \times \Theta_j = \Theta$ have utility over two possible *decisions*

$d \in \{d_0, d_1\}$. A decision d_0 yields payoff zero for any types. The payoff from a decision d_1 depends on both types θ. Hence, the payoff of agent i is $u_i(d,\theta) = v_i(\theta)\mathbf{1}\{d = d_1\}$. Note that $\{\theta \in \Theta | v_i(\theta) = 0\}$ is a subset of Θ for which agent i is indifferent between d_0 and d_1. A *decision rule* $\phi : \Theta \to [0,1]$ associates any type θ with *an outcome* $\phi(\theta)$, which is the probability of d_1. The joint distribution function $F(\theta)$ and its density function $f(\theta)$ are common knowledge. We assume $\Theta_i \times \Theta_j = [\underline{\theta}_i, \overline{\theta}_i] \times [\underline{\theta}_j, \overline{\theta}_j] \subset \mathbb{R}^2$.

Assumption 1: (a) *$v_i(\theta)$ and $v_j(\theta)$ are continuous and strictly increasing in both arguments.* (b) *The set $\{\theta \in \Theta | v_i(\theta) = v_j(\theta) = 0\}$ has at most finite number of elements.*

Assumption 2: (a) *$f(\theta)$ is continuous and strictly positive on Θ. The conditional density $f_i(\theta_j|\theta_i)$ is strictly positive on Θ_j.* (b) *For any subinterval $\widehat{\Theta}_j \subset \Theta_j$, the conditional distribution $F_i\left(\theta_j|\theta_i, \theta_j \in \widehat{\Theta}_j\right)$ is increasing in θ_i in the sense of first order stochastic dominance. The same condition applies by changing the role of i and j.*

Assumption 1 is illustrated in **Figure 1**.

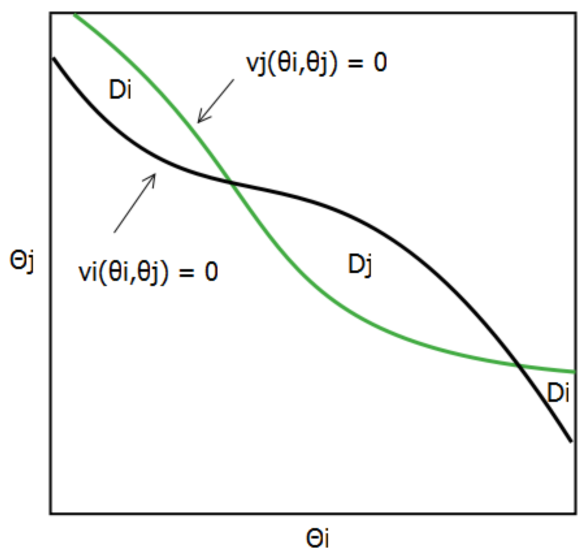

Figure 1. Indifference curves.

We measure θ_i in the horizontal direction and θ_j in the vertical direction so that "a vertical segment" means $\{\theta \in \Theta | \theta_i = a, \theta_j \in [b,c]\}$, while "a horizontal segment" means $\{\theta \in \Theta | \theta_i \in [a,b], \theta_j = c\}$. By Assumption 1, two indifference curves $v_i(\theta) = 0$ and $v_j(\theta) = 0$

are strictly decreasing and intersect at most finite number of times. Both agents prefer d_1 to d_0 in the area above two curves, while both prefer d_0 to d_1 in the area below two curves. Hence they agree in these areas. Agents disagree in the area labeled as D_i and D_j. In D_i, agent i prefers d_1 to d_0 while agent j prefers d_0 to d_1. The opposite is true in D_j.[2] Thus, the agreement/disagreement pattern may change across states under complete information. The efficient decision rule under complete information chooses any random mixture of two decisions in the area strictly between two curves and at their intersections, while it chooses d_1 (d_0) in the area above (below) the higher (lower) of the two curves.

Assumption 2 introduces a correlation of types, which, together with Assumption 1, makes $\int v_i(\theta_i,\theta_j)dF_i\left(\theta_j|\theta_i,\theta_j\in\widehat{\Theta}_j\right)$ monotonic in θ_i for any subinterval $\widehat{\Theta}_j\subset\Theta_j$. This is the expected payoff from d_1 for agent i of type θ_i conditional on beliefs that agent j's type lies in $\widehat{\Theta}_j$. Because of the monotonicity, the set $\left\{\theta_i\in\Theta_i|\int v_i\left(\theta_i,\theta_j\right)dF_i\left(\theta_j|\theta_i,\theta_j\in[x_{j1},x_{j2}]\right)=0\right\}$ has only one element if it is not empty. We use the following two functions for our analysis. For any subintervals $[x_{i1},x_{i2}]\subset\Theta_i$ and $[x_{j1},x_{j2}]\subset\Theta_j$,

$$\begin{aligned} h_i\left(x_{j1},x_{j2}\right) &\equiv \left\{\theta_i\in\Theta_i|\int v_i\left(\theta_i,\theta_j\right)dF_i\left(\theta_j|\theta_i,\theta_j\in[x_{j1},x_{j2}]\right)=0\right\} \\ &\quad \text{if the set is not empty, or} \\ &= \begin{cases} \underline{\theta}_i \text{ if } \int v_i\left(\theta_i,\theta_j\right)dF_i\left(\theta_j|\theta_i,\theta_j\in[x_{j1},x_{j2}]\right)>0\ \forall\theta_i\in\Theta_i \\ \overline{\theta}_i \text{ if } \int v_i\left(\theta_i,\theta_j\right)dF_i\left(\theta_j|\theta_i,\theta_j\in[x_{j1},x_{j2}]\right)<0\ \forall\theta_i\in\Theta_i \end{cases}, \end{aligned}$$

$$\begin{aligned} h_j\left(x_{i1},x_{i2}\right) &\equiv \left\{\theta_j\in\Theta_j|\int v_j\left(\theta_i,\theta_j\right)dF_j\left(\theta_i|\theta_j,\theta_i\in[x_{i1},x_{i2}]\right)=0\right\} \\ &\quad \text{if the set is not empty, or} \\ &= \begin{cases} \underline{\theta}_j \text{ if } \int v_j\left(\theta_i,\theta_j\right)dF_j\left(\theta_i|\theta_j,\theta_i\in[x_{i1},x_{i2}]\right)>0\ \forall\theta_j\in\Theta_j \\ \overline{\theta}_j \text{ if } \int v_j\left(\theta_i,\theta_j\right)dF_j\left(\theta_i|\theta_j,\theta_i\in[x_{i1},x_{i2}]\right)<0\ \forall\theta_j\in\Theta_j \end{cases}. \end{aligned}$$

[2] More precisely, they disagree in the area strictly inside two curves, both are indifferent at the intersections of two curves, only one agent is indifferent at points on the either curve exclusing the intersections, and both agree in the remaining area.

The function h_i determines an indifference type of i conditional on beliefs that agent j's type lies in $[x_{j1}, x_{j2}]$. Two functions h_i and h_j are strictly decreasing whenever they take the values other than $\underline{\theta}_i$, $\overline{\theta}_i$, $\underline{\theta}_j$, $\overline{\theta}_j$. We strengthen Green and Laffont's assumption with the following smoothness condition.[3]

Assumption 3: *h_i and h_j are continuous.*

A *mechanism* (M, g) is a pair of action space $M = M_i \times M_j$ and a measurable function $g : M \to [0, 1]$, where $g(m)$ is a probability of d_1 when $m = (m_i, m_j) \in M$ is chosen by agents. A *strategy* of agent i is a collection of conditional distributions $s_i(m_i|\theta_i)$, $\theta_i \in \Theta_i$. A pair of strategies $s = (s_i, s_j)$ in the mechanism (M, g) results in the decision rule $\phi(\theta) = \int_M g(m) ds_i(m_i|\theta_i) ds_j(m_j|\theta_j)$.

Definition 1: *A pair of strategies $s = (s_i, s_j)$ is a Bayes-Nash equilibrium of (M, g) if*

$$m_i^* \in \arg\max_{m_i \in M_i} \int_{\Theta_j} \int_{M_j} v_i(\theta_i, \theta_j) g(m) ds_j(m_j|\theta_j) dF_i(\theta_j|\theta_i)$$

for all m_i^* in the support of $s_i(m_i|\theta_i)$ for almost every θ_i, and

$$m_j^* \in \arg\max_{m_j \in M_j} \int_{\Theta_i} \int_{M_i} v_j(\theta_i, \theta_j) g(m) ds_i(m_i|\theta_i) dF_j(\theta_i|\theta_j)$$

for all m_j^* in the support of $s_j(m_j|\theta_j)$ for almost every θ_j.

Let $\mu^s(m, \theta)$ be the joint distribution over $M \times \Theta$ generated in a Bayes-Nash equilibrium s. Let $\mu_i^s(m_j|\theta_i)$ be the marginal distribution of m_j given θ_i. For every θ_i and $\mu_i^s(m_j|\theta_i)$-almost every m_j, define $F_i(\theta_j|\theta_i, m_j)$ to be the conditional distribution that agent i of type θ_i would hold about θ_j given m_j. Define $\mu_j^s(m_i|\theta_j)$ and $F_j(\theta_i|\theta_j, m_i)$ in a symmetric manner.

[3] In Green and Laffont, 1987, p.88, they applied Brouwer's fixed-point theorm to these functions. Therefore, it appears they also assumed this condition.

Definition 2: *A Bayes-Nash equilibrium s is a posterior equilibrium if*

$$\begin{aligned} m_i &\in \arg\max_{m_i' \in M_i} g(m_i', m_j) \int_{\Theta_j} v_i(\theta_i, \theta_j) dF_i(\theta_j | \theta_i, m_j), \\ m_j &\in \arg\max_{m_j' \in M_j} g(m_i, m_j') \int_{\Theta_i} v_j(\theta_i, \theta_j) dF_j(\theta_i | \theta_j, m_i), \end{aligned}$$

for $\mu^s(m, \theta)$-almost every (m, θ).

Definition 3: *A decision rule ϕ is posterior implementable if there is a mechanism (M, g) with a posterior equilibrium s which results in ϕ.*

Green and Laffont (1987) provide a complete characterization of a set of posterior implementable decision rules. Let H be the set of decreasing step functions ξ which partition $\Theta_i \times \Theta_j$ in two parts with the following properties:

$$\begin{aligned} &\text{any vertical segment } (\theta_i, [a, b]\,) \text{ of } \xi \text{ satisfies } \theta_i = h_i(a, b), \\ &\text{any horizontal segment } (\,[c, d]\,, \theta_j) \text{ of } \xi \text{ satisfies } \theta_j = h_j(c, d). \end{aligned} \tag{1}$$

Theorem (Green and Laffont, 1987)

Any posterior implementable decision rule ϕ is such that, for some $\xi \in H$, $\phi(\theta) = \phi^- \mathbf{1}\{\theta$ is below $\xi\} + \phi^+ \mathbf{1}\{\theta$ is above $\xi\}$ with $0 \leq \phi^- \leq \phi^+ \leq 1$.

In **Theorem**, the values of ϕ on ξ is ignored because the expected utility of agents is insensitive to it. **Figure 2** shows an example of posterior implementable decision rules.

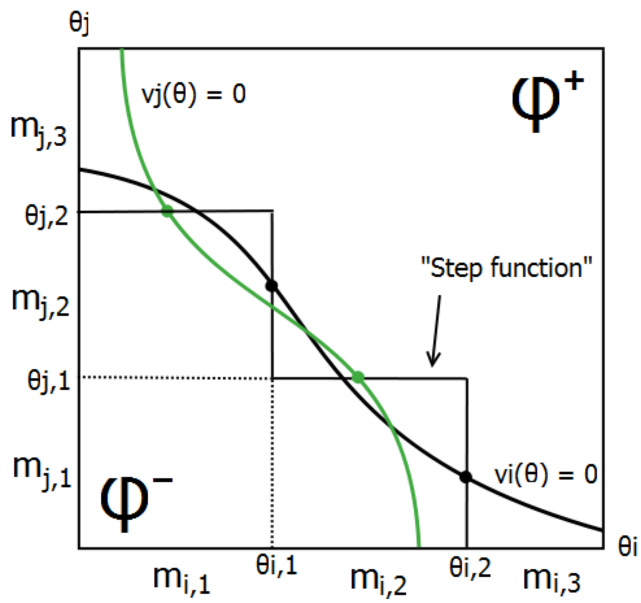

Figure 2. A posterior implementable decision rule with three actions for each agent.

In this example, each agent has three actions, each of which corresponds to a subinterval of types who uses that action in equilibrium. Each line segment in the step function has an associated indifferent type, characterized by one of the conditions (1) that defined the set H. For example, $\theta_{i,1}$ is a solution to $h_i(\theta_{j,1}, \theta_{j,2}) = \int_{\Theta_j} v_i(\theta_i, \theta_j) dF_i(\theta_j | \theta_i, \theta_j \in [\theta_{j,1}, \theta_{j,2}]) = 0$ so that agent i of type $\theta_{i,1}$ is indifferent between $m_{i,1}$ and $m_{i,2}$.

This characterization allows us to focus on a smaller class of mechanisms. If a mechanism (M, g) has a posterior equilibrium, there is an equivalent mechanism $\left(\widehat{M}, \widehat{g}\right)$ in terms of the resulting decision rule and the equilibrium information structure. An action space $\widehat{M}$ is *a set of subsets* of Θ. An action $\widehat{\Theta}_i \subset \Theta_i$ by agent i is interpreted as "my type is in $\widehat{\Theta}_i$", and it is truthful if his true type is in $\widehat{\Theta}_i$.[4] Moreover, we can assume without loss of generality that each action $\widehat{m}_i = \widehat{\Theta}_i$ is a closed interval and any two actions share at most one point.

For the rest of the paper, we use the following convention. Whenever a mechanism has multiple actions for one agent, we order them such that $\max \widehat{\Theta}_{i,k} = \min \widehat{\Theta}_{i,k+1}$ for kth and $k+1$th actions, and denote their boundary points by $\theta_{i,k} \equiv \max \widehat{\Theta}_{i,k} = \min \widehat{\Theta}_{i,k+1}$. If a mechanism has K actions available for agent i and L actions for agent j, we call it a (K, L)-

[4]For example, if $\Theta_i \in \widehat{M}$, "my type is in Θ_i" is truthful although it reveals no information.

mechanism, and the associated posterior implementable decision rule (if it exists) shall be called a (K, L)-*rule*. For any (K, L)-rules, K and L differ at most by one due to **Theorem** above. While the set of posterior implementable decision rules may look restrictive, K and L can be very large and $\phi^- \leq \phi^+$ can take any values in $[0, 1]$. It turns out that this environment is rich enough for the study of renegotiation-proofness.

3 Posterior Renegotiation-proof Decision Rules

Before formally defining PRP decision rules, we first investigate properties of posterior implementable decision rules. This helps us understand what the first mechanism can do to control information revelation before the second mechanism is proposed. The results will be used to characterize PRP decision rules defined in the following subsection.

3.1 Structure of posterior implementation

First, the following definitions will be useful for our purpose.

Definition 4: *A constant mechanism is* $M_i = \{\Theta_i\}$, $M_j = \{\Theta_j\}$, *and* $g(m) = \phi^0 \in [0, 1]$.

Definition 5: *A dictatorial mechanism for* i *is* $M_i = \left\{\widehat{\Theta}_{i,1}, \widehat{\Theta}_{i,2}\right\}$, $M_j = \{\Theta_j\}$ *and* $g(m) = \phi^- \mathbf{1}\left\{m_i = \widehat{\Theta}_{i,1}\right\} + \phi^+ \mathbf{1}\left\{m_i = \widehat{\Theta}_{i,2}\right\}$ *with* $0 \leq \phi^- < \phi^+ \leq 1$.

Definition 6: *A* $(2, 2)$*-mechanism is* $M_i = \left\{\widehat{\Theta}_{i,1}, \widehat{\Theta}_{i,2}\right\}$, $M_j = \left\{\widehat{\Theta}_{j,1}, \widehat{\Theta}_{j,2}\right\}$ *with* $0 \leq \phi^- < \phi^+ \leq 1$, *and it is either*

Low type: $g(m) = \phi^- \mathbf{1}\left\{m = \left(\widehat{\Theta}_{i,1}, \widehat{\Theta}_{j,1}\right)\right\} + \phi^+ \mathbf{1}\{otherwise\}$ *or*

High type: $g(m) = \phi^+ \mathbf{1}\left\{m = \left(\widehat{\Theta}_{i,2}, \widehat{\Theta}_{j,2}\right)\right\} + \phi^- \mathbf{1}\{otherwise\}$.

A constant mechanism always has a trivial posterior equilibrium and implements a $(1, 1)$-rule, but reveals no information.[5] The other two types of mechanisms may or may not have

[5] A constant mechanism is the only case in which, if a direct mechanism $M_i = \Theta_i$ is used instead of $M_i = \{\Theta_i\}$, every agent is indifferent among any actions after actions are made public. Hence, there is a posterior equilibrium in a pure strategy where type is perfectly revealed. However, once this happens, the efficient decision rule under complete information can be proposed. Anticipating this, the first mechanism

a posterior equilibrium. If a dictatorial mechanism for i has a posterior equilibrium, it implements a $(2,1)$-rule and partially reveals agent i's type but reveals nothing about agent j's type. If a $(2,2)$-mechanism has a posterior equilibrium, it implements a $(2,2)$-rule and it partially reveals information to both agents. If it is a low type, the low outcome ϕ^- needs low actions from both agents, while at least one high action implements ϕ^+. If it is a high type, the high outcome ϕ^+ needs high actions from both agents, while at least one low action implements ϕ^-. We say that a (K,K)-mechanism with $K \geq 2$ is a *low (high) type* if no agent can choose the low (high) outcome independent of the other agent's action. **Figure 2** shows a low type $(3,3)$-rule. By **Theorem**, any (K,K)-rule is either a low type or a high type. The following lemma studies the existence of posterior equilibrium.

Lemma 1 (existence of (K,L)-rules)

(i) *If no dictatorial mechanism for i (for j respectively) has a posterior equilibrium, then for all $K \geq 2$, no $(K+1,K)$-mechanism ($(K,K+1)$) has a posterior equilibrium.*

(ii) *If no $(2,2)$-mechanism has a posterior equilibrium, then for all $K \geq 3$, no (K,K)-mechanism has a posterior equilibrium.*

(iii) *A dictatorial mechanism for i can implement a $(2,1)$-rule if and only if*

$$\exists \theta_i^d \in \left(\underline{\theta}_i, \overline{\theta}_i\right) \ s.t.\ \theta_i^d = h_i\left(\underline{\theta}_j, \overline{\theta}_j\right).$$

(iv) *A $(2,2)$-mechanism of a low type can implement a $(2,2)$-rule if and only if*

$$\exists \theta' = \left(\theta_i', \theta_j'\right) \in \left(\underline{\theta}_i, \overline{\theta}_i\right) \times \left(\underline{\theta}_j, \overline{\theta}_j\right) \ s.t.\ \theta' = \left(h_i\left(\underline{\theta}_j, \theta_j'\right), h_j\left(\underline{\theta}_i, \theta_i'\right)\right).$$

(v) *A $(2,2)$-mechanism of a high type can implement a* $(2,2)$-rule *if and only if*

$$\exists \theta' = \left(\theta_i', \theta_j'\right) \in \left(\underline{\theta}_i, \overline{\theta}_i\right) \times \left(\underline{\theta}_j, \overline{\theta}_j\right) \ s.t.\ \theta' = \left(h_i\left(\theta_j', \overline{\theta}_j\right), h_j\left(\theta_i', \overline{\theta}_i\right)\right).$$

must be ex post incentive compatible, which is violated in general cases like **Figure 1**. We ignore this perfect information revelation case for the rest of the paper.

Lemma 1 (i) and **(ii)** show that dictatorial and $(2,2)$-mechanisms are the key mechanisms in terms of the existence of posterior equilibrium. Hence, for any information to be revealed in a posterior equilibrium, either dictatorial or $(2,2)$-mechanisms must have a posterior equilibrium. **Lemma 1 (iii)** through **(v)** present conditions for the existence of posterior equilibria in these mechanisms. The conditions state that there must be indifferent types that are consistent with an equilibrium information structure.

A $(2,1)$-rule represents a situation where a dictator i can optimally choose any outcome (ϕ^- or ϕ^+) based on his interim belief.[6] Agent j learns about the dictator's type through his choice, although she is forced to accept the dictator's choice and does not reveal any information. All $(K+1,K)$-rules share the property that only agent i can choose an outcome independent of the agent j's action. On the other hand, a $(2,2)$-rule give "equal rights" to agents because each agent can choose one particular outcome independent of what the other agent's action, while they need to cooperate in order to implement the other outcome.[7] Also, both agents learn new information in equilibrium. All (K,K)-rules share the property that both agents can choose the same outcome independent of the other agent's action, while they need to coordinate their actions to choose the other outcome.

Absence of dictators and reversed agreement. Because $(2,1)$, $(1,2)$, and $(2,2)$-rules are representative rules for all other posterior implementable rules, we investigate the conditions under which they exist (or do not exist) and their relationship in more detail. First, we study when information revelation in posterior equilibrium is not possible. By **Lemma 1 (i)(ii)**, it is impossible if and only if there is no $(2,1)$, $(1,2)$, and $(2,2)$-rule. Suppose no $(2,1)$ and $(1,2)$-rule exists. By **Lemma 1 (iii)** and its counterpart to $(1,2)$-

[6] Agent i can choose an outcome ϕ^- by choosing $\widehat{\Theta}_{i,1}$ and ϕ^+ by choosing $\widehat{\Theta}_{i,2}$.

[7] For a low type $(2,2)$-rule, agent i (j) can implement ϕ^+ by choosing $\widehat{\Theta}_{i,2}$ ($\widehat{\Theta}_{j,2}$) regardless of the other agent's action. To implement ϕ^-, coordinated actions $(\widehat{\Theta}_{i,1}, \widehat{\Theta}_{j,1})$ are required.

rules, there are four possible cases which we label as follows:

$$\begin{aligned} A(d_1) &: \left(h_i\left(\underline{\theta}_j, \overline{\theta}_j\right), h_j\left(\underline{\theta}_i, \overline{\theta}_i\right)\right) = \left(\underline{\theta}_i, \underline{\theta}_j\right), \\ A(d_0) &: \left(h_i\left(\underline{\theta}_j, \overline{\theta}_j\right), h_j\left(\underline{\theta}_i, \overline{\theta}_i\right)\right) = \left(\overline{\theta}_i, \overline{\theta}_j\right), \\ D(i) &: \left(h_i\left(\underline{\theta}_j, \overline{\theta}_j\right), h_j\left(\underline{\theta}_i, \overline{\theta}_i\right)\right) = \left(\underline{\theta}_i, \overline{\theta}_j\right), \\ D(j) &: \left(h_i\left(\underline{\theta}_j, \overline{\theta}_j\right), h_j\left(\underline{\theta}_i, \overline{\theta}_i\right)\right) = \left(\overline{\theta}_i, \underline{\theta}_j\right). \end{aligned}$$

The first case $A(d_1)$ is where two agents of all types agree that d_1 is better than d_0 given their interim beliefs $(\theta_i, \{\theta_j \in \Theta_j\})$ for agent i and $(\theta_j, \{\theta_i \in \Theta_i\})$ for agent j. We call it *interim agreement* on d_1. The second case $A(d_0)$ is where they agree that d_0 is better than d_1 and we call it interim agreement on d_0. The third case $D(i)$ is where agent i of all types prefers d_1 to d_0, while agent j of all types prefers d_0 to d_1. The fourth case $D(j)$ is the opposite disagreement pattern. We call $D(i)$ and $D(j)$ *interim disagreement*. The next result shows that, when the information revelation by dictatorial mechanisms is not possible, the interim agreement ($A(d_1)$ or $A(d_0)$) is necessary for the information revelation to occur.

Lemma 2 (Absence of dictators)

(i) *If there is interim disagreement, then there is no* $(2,2)$*-rule.*

(ii) *Suppose neither* $(1,2)$ *nor* $(2,1)$*-rule exists. If a* $(2,2)$*-rule of a low (high) type exists, then there is interim agreement on* d_1 *(on* d_0*).*

Lemma 2 (i) shows that interim disagreement is sufficient for no information revelation. On the other hand, when there is interim agreement, information may still be revealed by $(2,2)$-rules. **Lemma 2 (ii)** shows that, when coordinated actions can reveal information in the absence of dictators, coordinated actions must reverse the interim agreement. For example, a $(2,2)$-rule of *low* type exists only if there is interim agreement on the *high* decision d_1. In sum, no information revelation is possible if and only if (i) there is interim disagreement or (ii) there is interim agreement and no $(2,2)$-rule can reverse the agreement.

Figure 3 illustrates $A(d_1)$ with potential reversal, while **Figure 4** illustrates $A(d_1)$ with no possibility of reversal.

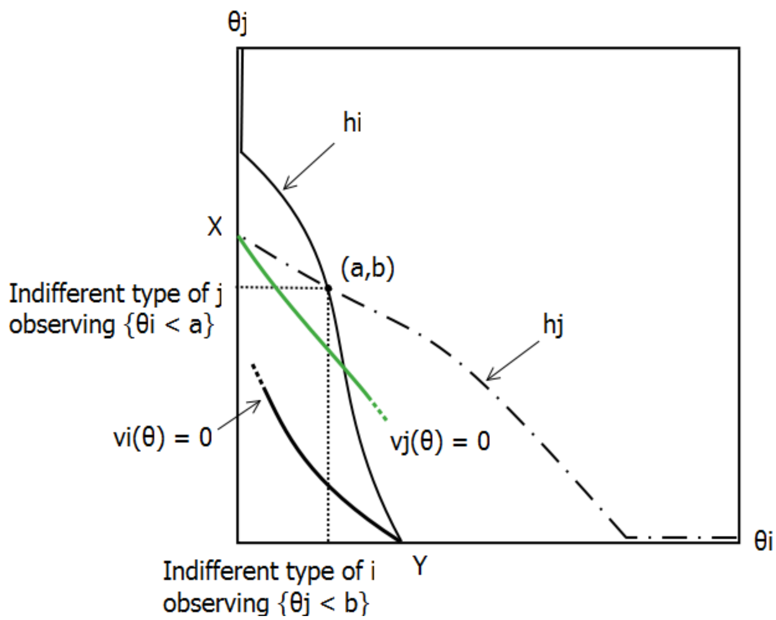

Figure 3. Interim agreement on d_1 and reversed agreement on d_0 on $\{\theta_i < a\} \times \{\theta_j < b\}$.

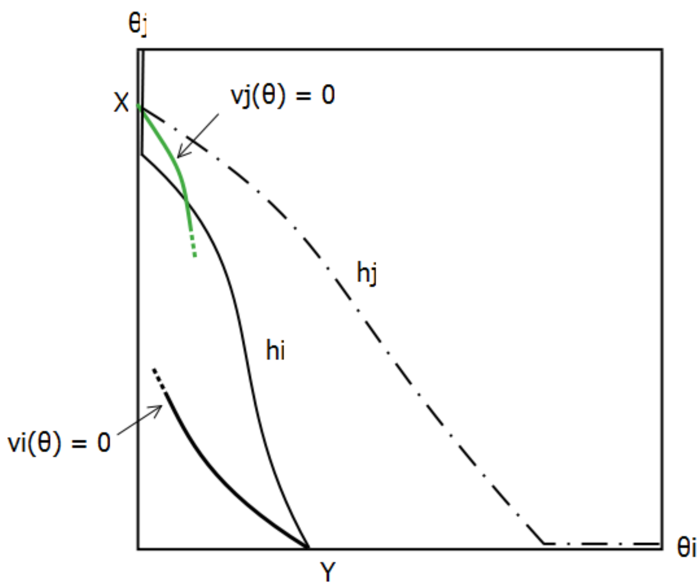

Figure 4. Interim agreement on d_1 and no reversal to d_0.

In **Figures 3** and **4**, a dashed line with the end point X is a function $h_j : \Theta_i \to \Theta_j$ defined by $h_j(\underline{\theta}_i, x_i)$ for $x_i \in \Theta_i$ and a solid line with the end point Y is a function $h_i : \Theta_j \to \Theta_i$ defined by $h_i(\underline{\theta}_j, x_j)$ for $x_j \in \Theta_j$. In both figures, the flat part of h_i implies that agent i

of $\underline{\theta}_i$ strictly prefers d_1 to d_0 given his interim belief. Similarly, the flat part of h_j implies that agent j of $\underline{\theta}_j$ strictly prefers d_1 to d_0 given his interim belief. Therefore, all agents of all types strictly prefer d_1 to d_0. This is interim agreement on d_1 labeled as $A(d_1)$. Given our assumptions, h_i and h_j are continuous and decreasing in the interior of Θ. Hence, h_i and h_j may or may not intersect in the interior of Θ.[8] But whey they do, the intersections of h_i and h_j represent $(2,2)$-rules of a low type by **Lemma 1 (iv)**. In **Figure 3**, conditional on $\{\theta_i < a\} \times \{\theta_j < b\}$, agents agree on d_0, not d_1. Thus, there is a $(2,2)$-rule that can reverse agreement. To see why information revelation must involve reversal here, consider agent i. According to his interim belief, agent i of all types prefers d_1 to d_0. However, conditional on the separation of agent j of lower types, his preference shifts towards d_0. Therefore, the separation of agent j of low type can provide incentive for agent i of lower types to separate, which in turn provides incentive for agent j of lower types to separate. However, the separation of agent j is possible only if there are high enough types who don't want to separate after observing the separation of low type agent i. If there were interim disagreement so that agent j of all types preferred d_0 to d_1, then the separation of agent i of low types would only *strengthen* the interim preference for all types of agent j. Thus, no separation would be possible if there is interim disagreement.

Dictators v.s. coordination. A natural question is whether a $(2,2)$-rule exists when both $(2,1)$- and $(1,2)$-rules exist. This is a situation where both agents may insist on his/her dictatorial mechanism, but someone (a third party) can suggest a $(2,2)$-rule as a compromising alternative. Such a suggestion not only put both parties on more equal footing but also facilitates more communication between them. When is such a suggestion possible? What is a nature of compromise? The next lemma uncovers some connections between dictatorial and $(2,2)$-rules.

Lemma 3 (dictatorial and $(2,2)$-rules)

Suppose both $(2,1)$- and $(1,2)$-rules exist. Let θ_i^d be the indifference type in the $(2,1)$-rule

[8] For example, h_i would be totally flat if $\{\theta \in \Theta | v_i(\theta) = 0\} = (\underline{\theta}_i, \underline{\theta}_j)$.

and θ_j^d be that in the $(1,2)$-rule.

(i) *If there are multiple ($k \geq 1$) $(2,2)$-rules of a low-type (respectively high), then k indifference points $\{(\theta_{i,1}^1, \theta_{j,1}^1), .., (\theta_{i,1}^k, \theta_{j,1}^k)\}$ lie in $(\theta_i^d, \overline{\theta}_i) \times (\theta_j^d, \overline{\theta}_j)$ ($(\underline{\theta}_i, \theta_i^d) \times (\underline{\theta}_j, \theta_j^d)$). If they are ordered by $\theta_{i,1}^1 < ... < \theta_{i,1}^k$, then $\theta_{j,1}^1 > ... > \theta_{j,1}^k$.*

Next, consider agent i with $\overline{\theta}_i$ (respectively $\underline{\theta}_i$) facing agent j with a type set $[\underline{\theta}_j, \theta_j^d]$ ($[\theta_j^d, \overline{\theta}_j]$), and agent j with $\overline{\theta}_j$ ($\underline{\theta}_j$) facing agent i with a type set $[\underline{\theta}_i, \theta_i^d]$ ($[\theta_i^d, \overline{\theta}_i]$).

(ii) *If both agents strictly prefer a higher outcome or both strictly prefer a lower outcome, then k is an odd number.*

(iii) *If one agent weakly prefers a higher outcome while the other agent weakly prefers a lower outcome, and preference is strict for at least one agent, then k is zero (non-existence) or an even number.*

(iv) *If both agents are indifferent for any outcomes, then k can be any number.*

Lemma 3 (i) shows a trade-off between a $(2,2)$-rule of a low type and a $(2,1)$-rule from agent i's point of view. In the $(2,2)$-rule, he cannot choose the low outcome ϕ^- unilaterally, which he could do in the $(2,1)$-rule. Also, while agent i can still implement ϕ^+ in the $(2,2)$-rule, the set of types who does so (e.g. $[\theta_i^k, \overline{\theta}_i]$ for the kth $(2,2)$-rule) shrinks compared with the set of such types in the $(2,1)$-rule ($[\theta_i^d, \overline{\theta}_i]$). The same applies to agent j if we compare the $(2,2)$-rule with a $(1,2)$-rule. Therefore, both agents make a compromise in the $(2,2)$-rule relative to his/her dictatorial rule. In return, however, both agents learn new information and are given a chance to coordinate their actions.

Lemma 3 also provides necessary conditions for the non-existence and uniqueness of $(2,2)$-rules given that either agent can be a dictator. If we rule out a non-generic case **(iv)**, conditions in **(ii)** and **(iii)** become necessary and sufficient. Hence, if no $(2,2)$-rule of a low-type exists, then it must be the case that the highest type agents facing the low-type dictator *disagree* on whether they prefer d_1 to d_0 or not. If there is a unique $(2,2)$-rule of a low-type, then they must agree on the matter. **Figure 5** illustrates a case where multiple $(2,2)$-rules of a low type exist, while **Figure 6** illustrates a case where they do not exist.

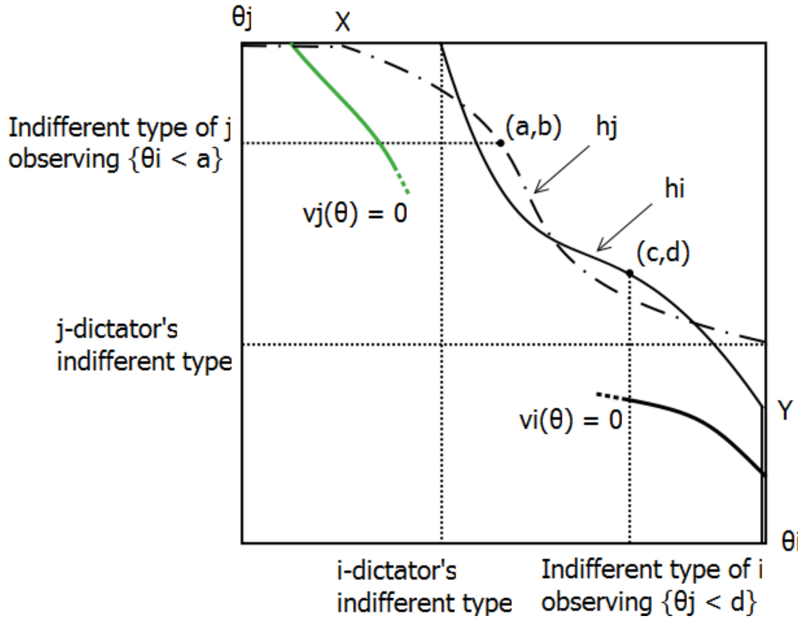

Figure 5. Co-existence of (2,2)-rules of a low type.

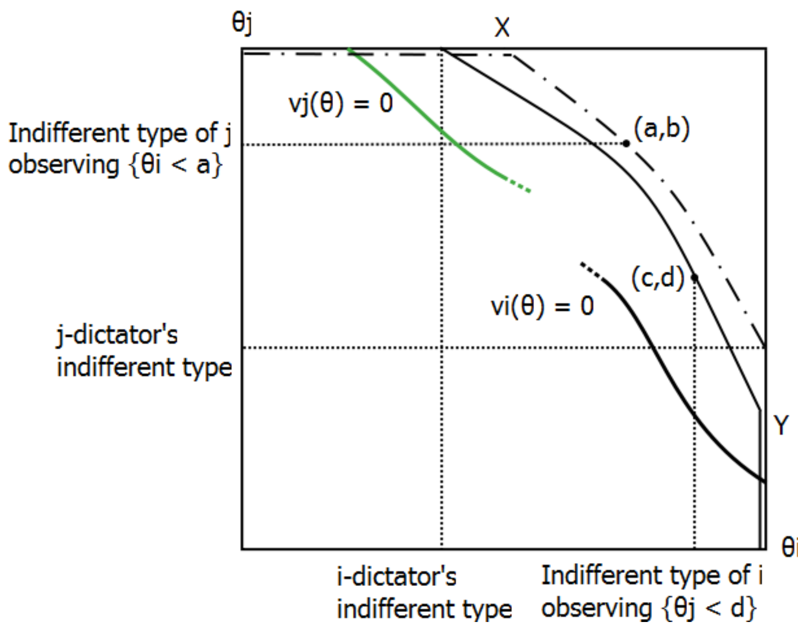

Figure 6. Non-existence of (2,2)-rules of a low type.

In **Figures 5** and **6**, a dashed line with the end point X is a function $h_j : \Theta_i \to \Theta_j$ defined by $h_j\left(\underline{\theta}_i, x_i\right)$ and a solid line with the end point Y is a function $h_i : \Theta_j \to \Theta_i$ defined by $h_i\left(\underline{\theta}_j, x_j\right)$. The intersections of h_i and h_j represent $(2, 2)$-rules of a low type. Points X and Y in **Figure 5** show that $\overline{\theta}_i$ facing $\widehat{\Theta}_{j,1}$ and $\overline{\theta}_j$ facing $\widehat{\Theta}_{i,1}$ both prefer d_1 to d_0 ("agreement"), while X and Y in **Figure 6** show that only $\overline{\theta}_i$ prefers d_1 to d_0 ("disagreement"). Note that the multiple $(2,2)$-rules exist when the two functions h_i and h_j have multiple intersections. When can we guarantee that a unique $(2,2)$-rule exists? The slope of h_i and h_j represents the *expected* marginal rate of substitution of types evaluated at the points of indifference, where expectation is taken over the other agent's type set with a different cutoff level. When two agents are sufficiently alike ex ante, a case in **Figure 5** is more likely. If additionally the expected marginal rate of substitution does not fluctuate much as the cutoff level changes (say h_i and h_j are close to linear), then there exists a unique $(2,2)$-rule. On the other hand, if two agents are sufficiently different ex ante, the case like **Figure 6** is possible, and a $(2,2)$-rule may not even exist.

The role of ex ante symmetry/asymmetry is formally stated below. Say agents are *symmetric* if $\Theta_i = \Theta_j = \left[\underline{\theta}, \overline{\theta}\right]$, $F_i(\cdot|a) = F_j(\cdot|a)$ $\forall a \in \left[\underline{\theta}, \overline{\theta}\right]$, and $v_i(a,b) = v_j(b,a)$ $\forall (a,b) \in \left[\underline{\theta}, \overline{\theta}\right]^2$. With symmetry, a $(2,1)$-rule exists if and only if $(1,2)$-rule exists. Let the dictator's indifference type be $\theta_i^d = \theta_j^d = \theta^d \in \left(\underline{\theta}, \overline{\theta}\right)$. Also, since h_i and h_j defined above are identical with symmetry, denote both by $h : \left[\underline{\theta}, \overline{\theta}\right] \to \left[\underline{\theta}, \overline{\theta}\right]$.

Lemma 4 (symmetric agents)

Suppose that agents are symmetric and a $(2,1)$-rule with the indifferent type θ^d exists.

(i) *A symmetric $(2,2)$-rule of a low type with $\theta_{i,1} = \theta_{j,1} = \theta^* \in \left(\theta^d, \overline{\theta}\right)$ exists.*

Next, consider agent with $\overline{\theta}$ facing the other agent with a type set $\left[\underline{\theta}, \theta^d\right]$.

(ii) *If this agent prefers d_0 to d_1, then:*

(a) *θ^* is the only $(2,2)$-rule of a low type if and only if $\forall \theta_i > \theta^*$, $h(\theta_i) < h^{-1}(\theta_i)$.*

(b) *$\theta^* \in \left(\frac{\theta^d + \overline{\theta}}{2}, \overline{\theta}\right)$ if h is concave.*

(iii) *If this agent prefers d_1 to d_0, then:*

(a) *θ^* is the only $(2,2)$-rule of a low type if and only if $\forall \theta_i < \theta^*, h(\theta_i) < h^{-1}(\theta_i)$.*

(b) *$\theta^* \in \left(\theta^d, \frac{\theta^d + \overline{\theta}}{2}\right)$ if h is convex.*

(iv) *If this agent is indifferent between d_1 and d_0, then:*

(a) *θ^* is the only $(2,2)$-rule of a low type if and only if $\forall \theta_i \in \left(\theta^*, \overline{\theta}\right)$ $h(\theta_i) \neq h^{-1}(\theta_i)$.*

(b) *$\theta^* \in \left(\theta^d, \frac{\theta^d + \overline{\theta}}{2}\right]$ $\left(\in \left[\frac{\theta^d + \overline{\theta}}{2}, \overline{\theta}\right)\right)$ if h is concave (convex).*

Suppose that agents are symmetric and $(2,1)$*-rule does not exists.*

(v) *A* $(2,2)$*-rule exists if and only if* $\{\theta \in \Theta | v_i(\theta) = 0\} \cap \left(\underline{\theta}, \overline{\theta}\right)^2$ *is not empty.*

An analogous result for a $(2,2)$-rule of a high type also holds but is not presented. From the converse of **Lemma 4 (i)**, given symmetric agents, if there is no symmetric $(2,2)$-rule, then there is neither $(1,2)$- nor $(2,1)$-rule. By **Lemma 1**, this implies that posterior implementable decision rules must be constant. Hence, with symmetry, information revelation without commitment to actions occurs if and only if $(2,2)$-rules exist. **Lemma 4 (v)** provides the condition for the absence of $(2,2)$-rules. With symmetry, information revelation without commitment to actions is impossible if and only if $v_i(\theta)$ (hence $v_j(\theta)$) does not go through the interior of Θ. This is the case where any agent of any type observing the other agent of any type prefer the same decision to the other decision.[9] Except this trivial situation, $(2,2)$-rules always exist if agents are symmetric.

3.2 PRP decision rules

We denote by $\Theta_\phi = \left\{\widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,l}\right\}_{k=1,..,K,l=1,..,L}$ a partition of Θ created by a (K,L)-rule. Because ϕ can realize only one of these $K \times L$ rectangle type sets, for each possible $\widehat{\Theta}_{k,l} = \widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,l} \in \Theta_\phi$, potential renegotiation is represented by posterior implementable decision rules defined on $\widehat{\Theta}_{k,l}$. For each $\widehat{\Theta}_{k,l} \in \Theta_\phi$, the relevant status quo outcome is $\phi\left(\widehat{\Theta}_{k,l}\right)$. Importantly, Assumptions 1 and 2 still hold if we replace Θ with any $\widehat{\Theta}_{k,l} \in \Theta_\phi$ in their statements.[10] Also, for each $\widehat{\Theta}_{k,l} \in \Theta_\phi$, Assumption 3 holds if we replace Θ_i with $\widehat{\Theta}_{i,k}$ and Θ_j with $\widehat{\Theta}_{j,l}$ in the definitions of h_i and h_j, given that it holds with Θ_i. Therefore, all the properties of posterior implementable rules defined on Θ apply to the posterior implementable rules defined on $\widehat{\Theta}_{k,l}$ by relabelling Θ with $\widehat{\Theta}_{k,l}$.

Definition 7: *Let* ϕ *be a* (K,L)*-rule with a partition* Θ_ϕ*. Say* ϕ *is NOT posterior*

[9] With the possible indifference for the lowest (highest) type observing the lowest (highest) type.

[10] In particular, note that Assumption 2 (b) was made for *any* subintervals $\widehat{\Theta}_j \subset \Theta_j$, which include any endogenously chosen subintervals.

renegotiation-proof (PRP) if $\exists \widehat{\Theta}_{k,l} \in \Theta_\phi$ *s.t. there is a decision rule* $\widehat{\phi}$ *defined on* $\widehat{\Theta}_{k,l}$ *with one of the following two properties:*

(a) $\widehat{\phi}$ *is posterior implementable and not constant, or*

(b) $\widehat{\phi}$ *is constant and both agents of all types in* $\widehat{\Theta}_{k,l}$ *weakly prefer* $\widehat{\phi}$ *to* $\phi\left(\widehat{\Theta}_{k,l}\right)$ *and at least one agent of some types strictly prefers* $\widehat{\phi}$ *to* $\phi\left(\widehat{\Theta}_{k,l}\right)$.

Posterior renegotiation-proofness is defined by the absence of both "posterior information revelation" and "unanimous improvement" *without* information revelation. In particular, the definition requires that there is no equilibrium in which information revelation occurs.[11] To the extent that information revelation *may* lead to some form of renegotiation, the definition above reflects our motivation to make the solution concept robust to details of the renegotiation process. Because **(b)** eliminates only unanimous improvement, there may be a constant rule that achieves improvement only for some types. Is there a renegotiation process that implements this constant rule only for the types that prefer it? If this is to be achieved, there must be information revelation. As long as such information revelation is subject to posterior implementability constraint, the whole process as a decision rule must be seen as a posterior implementable rule, and **(a)** takes care of the existence of such rule. Therefore, a constant rule that does not satisfy **(b)** may exist, but by definition of a PRP rule, it is impossible to render it into an improvement to a PRP rule.

In the above definition, the second mechanism that can challenge the first mechanism does not have to pass the same test applied to the first mechanism. We do not impose such a "credibility" restriction on $\widehat{\phi}$, because that would only weaken our solution concept and work against our motivation to make it robust. For example, if we consider competing firms proposing a mechanism as mentioned in the introduction, our notion of renegotiation-proofness seems reasonable in the following sense. Suppose that a firm finds a profitable opportunity (a mechanism that will be accepted by agents) but is aware that with some probability (i.e. for some types of agents) a competitor might steal its business later. We ar-

[11] This is similar to strong renegotiation-proofness in Maskin and Tirole (1992).

gue that this firm might take a chance and offer the mechanism anyway. In this environment, the definition above is meaningful because it provides sufficient robustness to competitors who do not (or fail to) care a "credibility" restriction.

The next proposition is the main result of the paper.

Proposition

(i) *If a* (K, L)*-rule is PRP, then* $K + L \leq 5$.

(ii) *A* $(1,1)$*-rule with* $\phi^0 \in [0,1]$ *is PRP if and only if one of the following conditions* (a)-(c) *holds:*

(a) *interim disagreement,*

(b) *interim agreement on* d_1, $\phi^0 = 1$, *and there is no* $(2,2)$*-rule of a low type,*

(c) *interim agreement on* d_0, $\phi^0 = 0$, *and there is no* $(2,2)$*-rule of a high type.*

(iii) *A* $(2,1)$*-rule with* $\theta_i^d \in (\underline{\theta}_i, \overline{\theta}_i)$ *and* $\phi^- < \phi^+$ *is PRP if and only if one of the following conditions* (a)-(c) *holds:*

(a) $h_j(\underline{\theta}_i, \theta_i^d) = \underline{\theta}_j$ *and* $\phi^+ = 1$,

(b) $h_j(\theta_i^d, \overline{\theta}_i) = \overline{\theta}_j$ *and* $\phi^- = 0$,

(c) $h_j(\underline{\theta}_i, \theta_i^d) = \overline{\theta}_j$, $h_j(\theta_i^d, \overline{\theta}_i) = \underline{\theta}_j$, $\phi^- = 0$, $\phi^+ = 1$. $(2,2)$*-rules of a high type on* $\widehat{\Theta}_{i,1} \times \Theta_j$ *and of a low type on* $\widehat{\Theta}_{i,2} \times \Theta_j$ *do not exist.*

(iv) *A* $(2,2)$*-rule of a low type with* $\phi^- < \phi^+$ *is PRP if and only if the following conditions are satisfied:*

(a) $(1,2)$*-rules and* $(2,2)$*-rules of a low type do not exist on* $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$.

(b) $(2,1)$*-rules and* $(2,2)$*-rules of a low type do not exist on* $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,2}$.

(c) $(2,2)$*-rules of a high type do not exist on* $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$.

(d) $\phi^- = 0$, $\phi^+ = 1$.

(v) (a) *A* $(3,2)$*-rule with* $\phi^- < \phi^+$ *is PRP if and only if* $\phi^- = 0$, $\phi^+ = 1$, *and* $(2,2)$*-rules of a low type on* $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,2}$ *and* $(2,2)$*-rules of a high type on* $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$ *do not exist.*

(b) *A* $(3,2)$*-rule is PRP only if* $\{\theta \in \Theta | v_j(\theta) = 0\} \subset \widehat{\Theta}_{i,2} \times \Theta_j$.

Proposition characterizes PRP decision rules.[12] A key feature of PRP decision rules is that they cannot have more than five actions. This is because whenever two agents have revealed that they are "middle" types, it is possible to make both agents reveal more information.[13] Considering that there is a continuum of types, and also that there can be a posterior implementable decision rules with infinite number of partitions[14], the result shows that information revelation is significantly reduced without commitment to a mechanism. Second, an intermediate outcome between zero and one (i.e. proposing random mixture of d_0 and d_1) can occur only in **(ii)(a)** where $\phi^0 \in [0,1]$ can take any value, or in **(iii)(a)** where $\phi^- < 1$ is possible, or in **(iii)(b)** where $\phi^+ > 0$ is possible. In **(ii)(a)**, there is interim disagreement. By **Lemma 2 (i)** constant rules are the only posterior implementable rules. Because there is interim disagreement, any change to ϕ^0 would be vetoed by one agent of all types without revealing any information. In **(iii)(a)**, there is interim disagreement conditional on $\left\{\theta_i < \theta_i^d\right\} \times \Theta_j$, where only agent j prefers d_1 to d_0. By **Lemma 2 (i)** applied on $\left\{\theta_i < \theta_i^d\right\} \times \Theta_j$, constant rules are the only posterior implementable rules. Due to interim disagreement conditional on the revealed information, any change to ϕ^- would be vetoed by one agent of all types without revealing any information. In **(iii)(b)**, the same explanation applies to disagreement created on $\left\{\theta_i^d < \theta_i\right\} \times \Theta_j$. These are the cases where a posterior equilibrium creates unambiguous disagreement independent of types.

For less trivial PRP rules **(iv)** and **(v)**, there is interim agreement conditional on each information partition created by the first mechanism. Thus, to avoid improvement by a constant rule, $\phi^- = 0$ and $\phi^+ = 1$ must hold. Figures below illustrate these rules.

[12] **(iii)** and **(v)** have symmetric counterparts for $(1,2)$- and $(2,3)$-rules but they are omitted.

[13] In fact, it is always possible to make both agents better off. See proof of (i).

[14] See Green and Laffont (1987) page 84 for a discussion of an accumulation point.

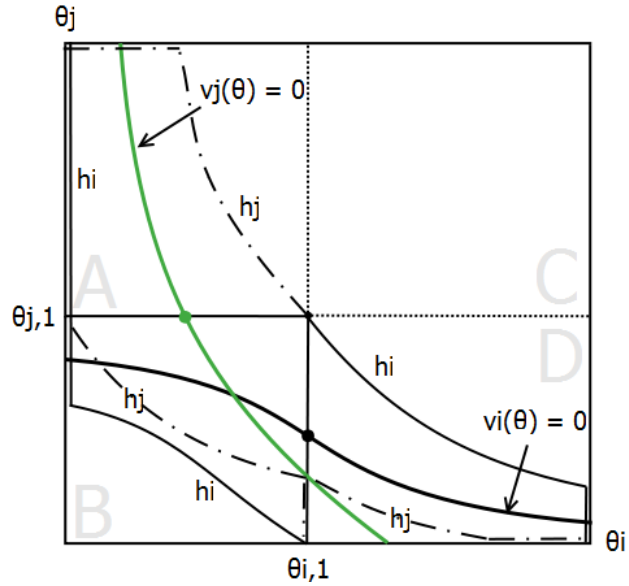

Figure 7. A posterior renegotiation-proof $(2,2)$-rule of low type.

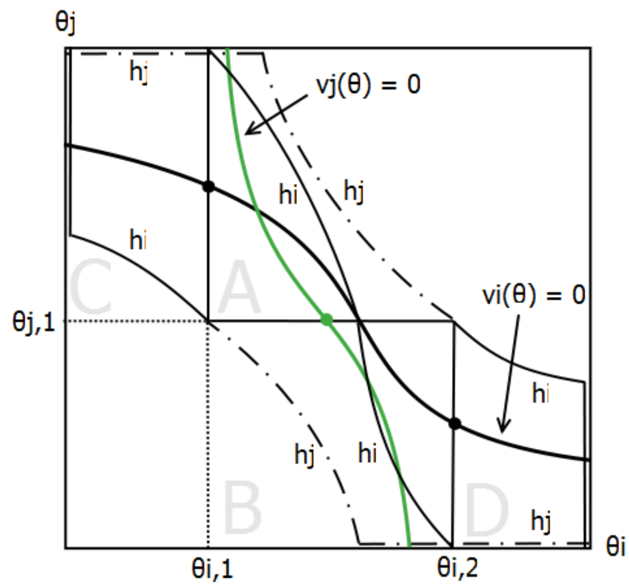

Figure 8. A posterior renegotiation-proof $(3,2)$-rule.

Figure 7 shows a PRP $(2,2)$-rule of low type. There are four partitions A, B, C, D created by this rule. The decision rule assigns the low outcome ϕ^- on B and the high outcome ϕ^+ on the other areas. Because $\theta_{j,1}$ is an indifference type given the belief $\{\theta_i < \theta_{i,1}\}$, the indifference curve $v_j(\theta) = 0$ must go through the interior of the boundary line between A and B. Similarly, $v_i(\theta) = 0$ must go through the interior of the boundary line between two B and D. On C, two agents' preferences are aligned in every state. If $\phi^+ < 1$, improvement with any higher outcome is possible on C. Thus, ϕ^+ must be one. On A, agents would disagree in the area left to $v_j(\theta) = 0$ (only agent i prefers d_1) and they agree on d_1 in the other area. However, conditional on the information revealed in the first mechanism, there is interim agreement, because agent j of all types $\theta_j \in [\theta_{j,1}, \overline{\theta}_j]$ weakly prefers d_1 given belief $\{\theta_i < \theta_{i,1}\}$. Moreover, no more information revelation is possible because there is neither $(1,2)$, $(2,1)$ nor $(2,2)$-rule on A.(No $(1,2)$-rule exists above $\theta_{j,1}$. Also, figure shows that $h_i(\theta_{j,1}, x_j)$ defined on $[\theta_{j,1}, \overline{\theta}_j]$ satisfies $h_i(\theta_{j,1}, x_j) = \underline{\theta}_i \ \forall x_j \in [\theta_{j,1}, \overline{\theta}_j]$. This implies there is neither $(2,1)$ nor $(2,2)$-rule.) Because $\phi^+ = 1$, the area A passes the test of PRP. On D, two agents agree on d_0 in the area below $v_j(\theta) = 0$, they agree on d_1 in the area above $v_i(\theta) = 0$, and they disagree in the area between two curves. However, conditional on the information revealed by the first mechanism, there is interim agreement on d_1, and no more information revelation is possible. With $\phi^+ = 1$, the area D also passes the test of PRP. Finally, on B, both agreement and disagreement are possible in different areas, and their patterns can change too. However, conditional on the information revealed by the first mechanism, there is interim agreement on d_0, and no more information revelation is possible. With $\phi^- = 0$, the area B also passes the test of PRP. All in all, the $(2,2)$-rule in **Figure 7** is robust to a posterior proposal of any posterior implementable rules. The rule itself looks very simple (two actions for each agent) and may appear to leave much room for potential renegotiation. For example, in the area between two indifference curves one agent would oppose to the mechanism's proposal, while there is the area in B and D where both agents would oppose to the mechanism's proposal. However, there is agreement conditional on the revealed information, because the potential disagreement or reversed agreement are outweighed in expectation. Moreover, it is common knowledge that those hidden states where disagreement or reversed agreement exist cannot be revealed without commitment to actions. We leave a similar analysis of **Figure 8** for interested readers.

Finally, when agents are ex ante symmetric, the characterization of PRP decision rules can be slightly simplified for $(1,1)$, $(2,1)$, $(1,2)$-rules, and an additional necessary condition can be added to $(2,2)$-rules.

Corollary (PRP with symmetric agents)

(i) *A $(1,1)$-rule with $\phi^0 \in [0,1]$ is PRP if and only if*

(a) $v_i(\theta) \geq 0 \ \forall \theta \in [\underline{\theta}, \overline{\theta}]$ *and* $\phi^0 = 1$, *or*

(b) $v_i(\theta) \leq 0 \ \forall \theta \in [\underline{\theta}, \overline{\theta}]$ *and* $\phi^0 = 0$.

(ii) *A $(2,1)$-rule with $\theta_i^d \in (\underline{\theta}, \overline{\theta})$ and $\phi^- < \phi^+$ is PRP if and only if* **Proposition (iii)(c)** *holds.*

(iii) *A $(2,2)$-rule with indifferent types $\theta_{i,1}$ and $\theta_{j,1}$ is PRP only if $\theta_{i,1} \neq \theta_{j,1}$.*

We comment only on the last result. By **Lemma 4 (v)**, with symmetry, a $(2,2)$-rule exists except the trivial case of no conflict of interests. However, if the rule is symmetric in the sense of $\theta_{i,1} = \theta_{j,1}$ (which always exists when any $(2,2)$-rule exists), then two of the information partitions $[\underline{\theta}, \theta_{i,1}] \times [\underline{\theta}, \theta_{j,1}]$ and $[\theta_{i,1}, \overline{\theta}] \times [\theta_{j,1}, \overline{\theta}]$ create symmetric environments again. Because v_i and v_j intersect in at least one of these two partitions, by **Lemma 4 (v)** applied to the new partition, there exists a $(2,2)$-rule on that partition. Thus, coordinated actions in a symmetric $(2,2)$-rule always invites renegotiation. With symmetric agents, in order to make a decision rule PRP, we have to make agents look asymmetric in equilibrium.

4 Conclusion

Surprisingly little is known about what we can achieve by using a mechanism in the presence of private information, once we leave the Revelation Principle and the implicit commitment assumptions behind it. This paper relaxes these assumptions and presents the concept of PRP decision rules. The results indicate that renegotiation-proofness puts significant restrictions on the amount of information revelation.

To further investigate the property of the proposed solution concept, it is important to obtain a characterization of posterior implementation in general environments. Also, the assumption that all actions are simultaneously chosen and made public constrains the equilibrium information structure. If any arbitrary observation pattern (including private observation) can be specified as a part of a mechanism, more information structures become possible in equilibrium. Allowing general patterns of information revelation will expand the set of implementable rules and will affect the set of PRP decision rules. Also, a continuum of types and costless renegotiation makes the set of PRP decision rules small in our environment. Studying how finite types and costly renegotiation can expand this set seems important for many applications. Finally, by allowing agents to adjust the value of (ϕ^-, ϕ^+), the model could be extended to study dynamic learning and coordination.[15] Relative to an exogenous proposal by a third party, a strategic proposal by informed agents is likely to be more constrained, hence likely to enlarge the set of renegotiation-proof rules.

5 Proofs

Proof of Lemma 1.

(i) If a $(K+1, K)$-mechanism has a posterior equilibrium, $\underline{\theta}_i < \theta_{i,1} = h_i\left(\theta_{j,K-1}, \overline{\theta}_j\right)$ and $\theta_{i,K} = h_i\left(\underline{\theta}_j, \theta_{j,1}\right) < \overline{\theta}_i$. Because h_i is strictly decreasing in both arguments, $\theta_{i,1} = h_i\left(\theta_{j,K-1}, \overline{\theta}_j\right) < h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) < h_i\left(\underline{\theta}_j, \theta_{j,1}\right) = \theta_{i,K}$. Because $\theta_i^d \equiv h_i\left(\underline{\theta}_j, \overline{\theta}_j\right)$ is an indifferent type of i given his prior belief, $\widehat{\Theta}_{i,1} = \left[\underline{\theta}_i, \theta_i^d\right]$ and $\widehat{\Theta}_{i,2} = \left[\theta_i^d, \overline{\theta}_i\right]$ has a posterior equilibrium. The proof is symmetric for a $(K, K+1)$-mechanism.

(ii) Consider a (K, K)-mechanism of a low type (as in **Figure 2**). If it has a posterior equilibrium, $\underline{\theta}_i < \theta_{i,1} = h_i\left(\theta_{j,K-2}, \theta_{j,K-1}\right)$ and $\theta_{i,K-1} = h_i\left(\underline{\theta}_j, \theta_{j,1}\right) < \overline{\theta}_i$. Similarly, both $\underline{\theta}_j < \theta_{j,1} = h_j\left(\theta_{i,K-2}, \theta_{i,K-1}\right)$ and $\theta_{j,K-1} = h_j\left(\underline{\theta}_i, \theta_{i,1}\right) < \overline{\theta}_j$. Let $X_i \equiv \bigcup_{k=2}^{K-1} \widehat{\Theta}_{i,k}$ and $X_j \equiv \bigcup_{k=2}^{K-1} \widehat{\Theta}_{j,k}$. Because h_i and h_j are strictly decreasing in both arguments, $\min X_i = \theta_{i,1} = h_i\left(\theta_{j,K-2}, \theta_{j,K-1}\right) < h_i\left(\underline{\theta}_j, \theta_{j,K-1}\right) < h_i\left(\underline{\theta}_j, \theta_{j,1}\right) = \theta_{i,K-1} = \max X_i$ and $\min X_j = \theta_{j,1} = h_j\left(\theta_{i,K-2}, \theta_{i,K-1}\right) < h_j\left(\underline{\theta}_i, \theta_{i,K-1}\right) < h_j\left(\underline{\theta}_i, \theta_{i,1}\right) = \theta_{j,K-1} = \max X_j$. Therefore, $\left(x_i^*, x_j^*\right) \equiv \left(h_i\left(\underline{\theta}_j, \theta_{j,K-1}\right), h_j\left(\underline{\theta}_i, \theta_{i,K-1}\right)\right)$ lies in the interior of $X_i \times X_j$. Define $h_i : X_j \to X_i$

[15] Watson (1999) is a related work in a dynamic environment.

by $h_i\left(\underline{\theta}_j, x_j\right)$ $\forall x_j \in X_j$. Similarly define $h_j : X_i \rightarrow X_j$ by $h_j(\underline{\theta}_i, x_i)$ $\forall x_i \in X_i$. Note that $h_i(\cdot)$ and $h_j(\cdot)$ are continuous, strictly decreasing, $h_i(\theta_{j,1}) = \theta_{i,K-1}$, $h_j(\theta_{i,1}) = \theta_{j,K-1}$, $h_i(\theta_{j,K-1}) = x_i^*$, and $h_j(\theta_{i,K-1}) = x_j^*$. Hence, a mapping $t : X_i \times X_j \rightarrow X_i \times X_j$ defined by $t(x_i, x_j) = (h_i(x_j), h_j(x_i))$ has at least one fixed point in $(x_i^*, \theta_{i,K-1}) \times \left(x_j^*, \theta_{j,K-1}\right)$. This fixed point constitutes a $(2,2)$-rule of low type. The proof is symmetric for a high type.

(iii) (if) By Assumptions 1 and 2, $\theta_i \lesseqgtr \theta_i' \Leftrightarrow \int_{\Theta_j} v_i(\theta_i, \theta_j) dF_i(\theta_j|\theta_i) \lesseqgtr 0$. Agent i with $\theta_i < \theta_i'$ prefers a higher outcome while agents with $\theta_i > \theta_i'$ prefers a lower outcome. Therefore, a mechanism with $\widehat{\Theta}_{i,1} = [\underline{\theta}_i, \theta_i']$ leading to ϕ^- and $\widehat{\Theta}_{i,2} = \left[\theta_i', \overline{\theta}_i\right]$ leading to $\phi^+ > \phi^-$ has a posterior equilibrium.

(only if) If the condition is not satisfied, it must be either

$$\begin{aligned} \forall \theta_i &\in \Theta_i, \int_{\Theta_j} v_i(\theta_i, \theta_j) dF_i(\theta_j|\theta_i) \geq 0 \Leftrightarrow h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \underline{\theta}_i \text{ or} \\ \forall \theta_i &\in \Theta_i, \int_{\Theta_j} v_i(\theta_i, \theta_j) dF_i(\theta_j|\theta_i) \leq 0 \Leftrightarrow h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \overline{\theta}_i. \end{aligned}$$

For the former case, agent i always prefers a higher outcome and $\widehat{\Theta}_{i,1}$ would not be chosen. For the latter case,agent i always prefers a lower outcome and $\widehat{\Theta}_{i,2}$ would not be chosen.

(iv) (if) For agent i observing $\theta_j \in \left[\underline{\theta}_j, \theta_j'\right]$, $\theta_i \lesseqgtr \theta_i' \Leftrightarrow \int_{\Theta_j} v_i(\theta_i, \theta_j) dF_i\left(\theta_j|\theta_i, \theta_j \in \left(\underline{\theta}_j, \theta_j'\right)\right) \lesseqgtr 0$. Hence, agent i with $\theta_i < \theta_i'$ prefers a lower outcome and optimally chooses $\widehat{\Theta}_{i,1}$ while agent i with $\theta_i > \theta_i'$ prefers a higher outcome and optimally chooses $\widehat{\Theta}_{i,2}$. For agent i observing $\theta_j \in \left[\theta_j', \overline{\theta}_j\right]$, the outcome ϕ^+ is independent of his action and choosing $\widehat{\Theta}_{i,1}$ or $\widehat{\Theta}_{i,2}$ truthfully is optimal.

(only if) If the condition is not satisfied, it must be one of the six cases:

(a) $\forall \theta_i \in \Theta_i$, $h_i\left(\underline{\theta}_j, x_j\right) = \underline{\theta}_i$ for all $x_j \in \left(\underline{\theta}_j, \overline{\theta}_j\right)$,

(b) $\forall \theta_i \in \Theta_i$, $h_i\left(\underline{\theta}_j, x_j\right) = \overline{\theta}_i$ for all $x_j \in \left(\underline{\theta}_j, \overline{\theta}_j\right)$,

(c) $\forall x_j \in \left(\underline{\theta}_j, \overline{\theta}_j\right)$ s.t. $h_i\left(\underline{\theta}_j, x_j\right) \in \left(\underline{\theta}_i, \overline{\theta}_i\right)$, it holds

$$\int_{\Theta_i} v_j(\theta_i, x_j) dF_j\left(\theta_i | x_j, \theta_i \in \left[\underline{\theta}_i, h_i\left(\underline{\theta}_j, x_j\right)\right]\right) = c \neq 0.$$

Cases (d), (e), (f) are symmetric cases for agent j. For (a) and (b), agent i always prefers one outcome to the other (ϕ^+ for (a) and ϕ^- for (b)), so there cannot be a $(2,2)$-posterior equilibrium. For (c), first note that $h_i\left(\underline{\theta}_j, x_j\right)$ is a candidate for $\theta_{i,1}$. If $c > 0$, then $h_j\left(\underline{\theta}_i, h_i\left(\underline{\theta}_j, x_j\right)\right) < x_j$. All types in $\left(h_j\left(\underline{\theta}_i, h_i\left(\underline{\theta}_j, x_j\right)\right), x_j\right]$ would prefer a higher outcome and hence choose $\widehat{\Theta}_{j,2} = \left[x_j, \overline{\theta}_j\right]$ rather than $\widehat{\Theta}_{j,1} = \left[\underline{\theta}_j, x_j\right]$ as they are supposed to. If $c < 0$, then $h_j\left(\underline{\theta}_i, h_i\left(\underline{\theta}_j, x_j\right)\right) > x_j$. All types in $\left[x_j, h_j\left(\underline{\theta}_i, h_i\left(\underline{\theta}_j, x_j\right)\right)\right)$ prefer a lower outcome and hence choose $\widehat{\Theta}_{j,1}$ rather than $\widehat{\Theta}_{j,2}$. The remaining cases are symmetric.

(v) Symmetric with (iv). ■

Proof of Lemma 2.

First, we show that neither $(2,1)$ nor $(1,2)$-rule exists if and only if one of $A(d_0)$, $A(d_1)$, $D(i)$, $D(j)$ is true. By Lemma 1 (iii), $(2,1)$-rules do not exist if and only if $h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \underline{\theta}_i$ or $h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \overline{\theta}_i$. Similarly, $(2,1)$-rules do not exist if and only if $h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \underline{\theta}_j$ or $h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \overline{\theta}_j$. Hence, there are four possible combinations. First, $h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \underline{\theta}_i$ and $h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \underline{\theta}_j$ imply both agents of all types prefer d_1 to d_0 ($A(d_1)$), while $h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \overline{\theta}_i$ and $h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \overline{\theta}_j$ imply both agents of all types prefer d_0 to d_1 ($A(d_0)$). Next, $h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \underline{\theta}_i$ and $h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \overline{\theta}_j$ imply agent i of all types prefer d_1 to d_0, while agent j of all types prefer d_0 to d_1 ($D(i)$). Finally, $h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \overline{\theta}_i$ and $h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \underline{\theta}_j$ imply agent i of all types prefer d_0 to d_1, while agent j of all types prefer d_1 to d_0 ($D(j)$).

(i) Consider $D(i)$. $(2,2)$-rules of high type cannot exist because $h_i\left(x_j, \overline{\theta}_j\right) \leq h_i\left(\underline{\theta}_j, \overline{\theta}_j\right) = \underline{\theta}_i$ $\forall x_j \in \Theta_j$. $(2,2)$-rules of low type cannot exist because $h_j\left(\underline{\theta}_i, x_i\right) \geq h_j\left(\underline{\theta}_i, \overline{\theta}_i\right) = \overline{\theta}_j$ $\forall x_i \in \Theta_i$. The case $D(j)$ is similar.

(ii) Suppose that neither $(2,1)$- nor $(1,2)$-rule exists and also that a $(2,2)$-rule of a low type exists. By (i), there is interim agreement. Suppose it is $A(d_0)$. Let $\theta_{i,1}$ and $\theta_{j,1}$ be indifferent types for a $(2,2)$-rule of a low type. Then $\theta_{i,1} = h_i\left(\underline{\theta}_j, \theta_{j,1}\right) < \overline{\theta}_i = h_i\left(\underline{\theta}_j, \overline{\theta}_j\right)$. This contradicts that h_i is decreasing in the second argument. Thus, the interim agreement must be $A(d_1)$. ■

Proof of Lemma 3.

(i) Define $h_i : \Theta_j \to \Theta_i$ by $h_i\left(\underline{\theta}_j, x_j\right)$ and $h_j : \Theta_i \to \Theta_j$ by $h_j\left(\underline{\theta}_i, x_i\right)$. Then $h_i\left(\overline{\theta}_j\right) = \theta_i^d$ and h_i is strictly decreasing. Similarly, $h_j\left(\overline{\theta}_i\right) = \theta_j^d$ and h_j is strictly decreasing. By **Lemma 1**, $(2,2)$-rules of a low type exist if and only if h_i and h_j have intersections. By construction of h_i and h_j, all the intersections must lie in $\left(\theta_i^d, \overline{\theta}_i\right) \times \left(\theta_j^d, \overline{\theta}_j\right)$ and satisfy the last property.

(ii) The stated condition implies either (a) $h_i\left(\theta_j^d\right) < \overline{\theta}_i$ and $h_j\left(\theta_i^d\right) < \overline{\theta}_j$ (this is **Figure 5**) or (b) $\exists \theta_i > \theta_i^d$ s.t. $h_j\left(\theta_i\right) = \overline{\theta}_j$ and $\exists \theta_j > \theta_j^d$ s.t. $h_i\left(\theta_j\right) = \overline{\theta}_i$. Either way, because h_i and h_j are strictly decreasing, they must cross each other for odd number of times in $\left[\theta_i^d, \overline{\theta}_i\right] \times \left[\theta_j^d, \overline{\theta}_j\right]$.

(iii) The stated condition implies either (a) $h_i\left(\theta_j^d\right) \leq \overline{\theta}_i$ and $\exists \theta_i \geq \theta_i^d$ s.t. $h_j\left(\theta_i\right) = \overline{\theta}_j$ and at least one inequality is strict, or (b) $h_j\left(\theta_i^d\right) \leq \overline{\theta}_j$ and $\exists \theta_j \geq \theta_j^d$ s.t. $h_i\left(\theta_j\right) = \overline{\theta}_i$ and at least one inequality is strict. Either way, h_i and h_j must cross each other for even number of times in $\left[\theta_i^d, \overline{\theta}_i\right] \times \left[\theta_j^d, \overline{\theta}_j\right]$.

(iv) The stated condition implies that h_i and h_j intersect at two points $\left(\theta_i^d, \overline{\theta}_j\right)$ and $\left(\overline{\theta}_i, \theta_j^d\right)$, but these two points do not constitute $(2,2)$-rules. Two strictly decreasing curves sharing the same end points can intersect in the middle any number of times. ■

Proof of Lemma 4.

(i) When agents are symmetric, case **(iii)** in **Lemma 3** cannot happen because h_i and h_j are located symmetrically with respect to a straight line connecting two points $(\underline{\theta}, \underline{\theta})$ and $\left(\overline{\theta}, \overline{\theta}\right)$ on $\left[\underline{\theta}, \overline{\theta}\right]^2$. Call this line the 45 degree line. Because h_i and h_j are strictly decreasing and cross the 45 degree line, they must cross each other on the 45 degree line only once. This is θ^*, and by **Lemma 3 (i)**, $\theta^* > \theta^d$.

(ii) (a) $h_i\left(\theta_j\right)$ starts from the interior of a segment $\left\{\overline{\theta}, \left[\theta^d, \overline{\theta}\right]\right\}$ and monotonically decreases in θ_j to a point $\left(\theta^d, \overline{\theta}\right)$, while $h_j\left(\theta_i\right)$ starts from the interior of a segment $\left\{\left[\theta^d, \overline{\theta}\right], \overline{\theta}\right\}$ and monotonically decreases in θ_i to a point $\left(\overline{\theta}, \theta^d\right)$. If the condition $\forall \theta_i > \theta^*\ h\left(\theta_i\right) < h^{-1}\left(\theta_i\right)$ holds, h_i and h_j do not cross below the 45 degree line. By symmetry, θ^* is the only intersection. If the condition does not hold, h_i and h_j have an intersection below and above the 45 degree line and θ^* is not a unique $(2,2)$-rule.

(b) If h is concave, both h_i and h_j must lie above the straight line connecting two points $\left(\theta^d, \overline{\theta}\right)$ and $\left(\overline{\theta}, \theta^d\right)$. Call this line the negative 45 degree line. Because the point $\left(\frac{\theta^d+\overline{\theta}}{2}, \frac{\theta^d+\overline{\theta}}{2}\right)$ is on the negative 45 line, $\theta^* \in \left(\frac{\theta^d+\overline{\theta}}{2}, \overline{\theta}\right)$.

(iii) $h_i(\theta_j)$ starts from the interior of a segment $\left\{\overline{\theta}, \left[\underline{\theta}, \theta^d\right]\right\}$ and decreases in θ_j to a point $\left(\theta^d, \overline{\theta}\right)$, while $h_j(\theta_i)$ starts from the interior of a segment $\left\{\left[\underline{\theta}, \theta^d\right], \overline{\theta}\right\}$ and decreases in θ_i to a point $\left(\overline{\theta}, \theta^d\right)$. The rest of the proof is analogous to (ii) and hence omitted.

(iv) $h_i(\theta_j)$ starts from a point $\left(\overline{\theta}, \theta^d\right)$ and decreases in θ_j to a point $\left(\theta^d, \overline{\theta}\right)$, while $h_j(\theta_i)$ starts from a point $\left(\theta^d, \overline{\theta}\right)$ and decreases in θ_i to a point $\left(\overline{\theta}, \theta^d\right)$. The rest of the proof is analogous to (ii) and hence omitted.

(v) If $\{\theta \in \Theta | v_i(\theta) = 0\} \cap \left(\underline{\theta}, \overline{\theta}\right)^2$ is empty, either $\{\theta \in \Theta | v_i(\theta) = 0\} = (\underline{\theta}, \underline{\theta})$ or $\{\theta \in \Theta | v_i(\theta) = 0\} = \left(\overline{\theta}, \overline{\theta}\right)$. For the former, $h_i(\underline{\theta}, x_j) = \underline{\theta}\ \forall x_j \in \Theta_j$ while for the latter $h_i\left(x_j, \overline{\theta}\right) = \overline{\theta}\ \forall x_j \in \Theta_j$. Hence, there is no $(2,2)$-rule. If a $(2,2)$-rule exists, $v_i(\theta) = 0$ must go through the interior of vertical segment $\{\theta \in \Theta | \theta_i = \theta_{i,1}, \theta_j \in [\underline{\theta}, \theta_{j,1}]\}$ or $\left\{\theta \in \Theta | \theta_i = \theta_{i,1}, \theta_j \in \left[\theta_{j,1}, \overline{\theta}\right]\right\}$, because $\theta_{i,1}$ satisfies either $\theta_{i,1} = h_i(\underline{\theta}, \theta_{j,1})$ or $\theta_{i,1} = h_i\left(\theta_{j,1}, \overline{\theta}\right)$. ■

Proof of Proposition.

(i) Consider a (K, L)-rule with $K + L \geq 6$. At least one rectangle $\widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,l}$ on which the high outcome ϕ^+ is chosen is characterized by

$$\begin{aligned}
\underline{\theta}'_i &\equiv \min \widehat{\Theta}_{i,k} = h_i(\theta_{j,l-1}, \theta_{j,l}) \text{ and } \overline{\theta}'_i \equiv \max \widehat{\Theta}_{i,k} = h_i(\theta_{j,l-2}, \theta_{j,l-1}), \\
\underline{\theta}'_j &\equiv \min \widehat{\Theta}_{j,l} = h_j(\theta_{i,k-1}, \theta_{i,k}) \text{ and } \overline{\theta}'_j \equiv \max \widehat{\Theta}_{j,l} = h_j(\theta_{i,k-2}, \theta_{i,k-1}).
\end{aligned}$$

Define $h_i : \widehat{\Theta}_{j,l} \to \widehat{\Theta}_{i,k}$ by

$$\begin{aligned} h_i\left(\underline{\theta}'_j, x_j\right) &\equiv \left\{\theta_i \in \widehat{\Theta}_{i,k} | \int v_i(\theta_i, \theta_j) dF_i(\theta_j|\theta_i, \theta_j \in \left[\underline{\theta}'_j, x_j\right]) = 0\right\} \\ &\quad \text{if the set is not empty,} \\ &= \begin{cases} \underline{\theta}'_i \text{ if } \int v_i\left(\theta_i, \theta_j\right) dF_i\left(\theta_j|\theta_i, \theta_j \in \left[\underline{\theta}'_j, x_j\right]\right) > 0 \ \forall \theta_i \in \widehat{\Theta}_{i,k} \\ \overline{\theta}'_i \text{ if } \int v_i\left(\theta_i, \theta_j\right) dF_i\left(\theta_j|\theta_i, \theta_j \in \left[\underline{\theta}'_j, x_j\right]\right) < 0 \ \forall \theta_i \in \widehat{\Theta}_{i,k} \end{cases} \end{aligned}$$

for $x_j \in \widehat{\Theta}_{j,l}$. Similarly define $h_j : \widehat{\Theta}_{i,k} \to \widehat{\Theta}_{j,l}$ by

$$\begin{aligned} h_j(\underline{\theta}'_i, x_i) &\equiv \left\{\theta_j \in \widehat{\Theta}_{j,l} | \int v_j(\theta_i, \theta_j) dF_j(\theta_i|\theta_j, \theta_i \in [\underline{\theta}'_i, x_i]) = 0\right\} \\ &\quad \text{if the set is not empty,} \\ &= \begin{cases} \underline{\theta}'_j \text{ if } \int v_j\left(\theta_i, \theta_j\right) dF_j\left(\theta_i|\theta_j, \theta_i \in [\underline{\theta}'_i, x_i]\right) > 0 \ \forall \theta_j \in \widehat{\Theta}_{j,l} \\ \overline{\theta}'_j \text{ if } \int v_j\left(\theta_i, \theta_j\right) dF_j\left(\theta_i|\theta_j, \theta_i \in [\underline{\theta}'_i, x_i]\right) < 0 \ \forall \theta_j \in \widehat{\Theta}_{j,l} \end{cases} \end{aligned}$$

for $x_i \in \widehat{\Theta}_{i,k}$. Note that $h_i(\cdot)$ and $h_j(\cdot)$ are continuous, strictly decreasing in the interior of $\widehat{\Theta}_{k,l}$, $h_i\left(\underline{\theta}'_j, \overline{\theta}'_j\right) = h_i\left(\theta_{j,l-1}, \theta_{j,l}\right) = \underline{\theta}'_i$ and $h_j(\underline{\theta}'_i, \overline{\theta}'_i) = h_j\left(\theta_{i,k-1}, \theta_{i,k}\right) = \underline{\theta}'_j$. Also, $h_i\left(\underline{\theta}'_j, \underline{\theta}'_j\right) = \left\{\theta_i \in \Theta_i | v_i\left(\theta_i, \underline{\theta}'_j\right) = 0\right\} \in \left(\underline{\theta}'_i, \overline{\theta}'_i\right)$, i.e., $v_i(\theta) = 0$ goes through a horizontal segment $\left\{\theta | \theta_i \in \widehat{\Theta}_{i,k}, \theta_j = \underline{\theta}'_j\right\}$, because $v_i(\theta) = 0$ goes through two vertical segments $\left\{\theta | \theta_i = \underline{\theta}'_i, \theta_j \in \widehat{\Theta}_{j,l}\right\}$ and $\left\{\theta | \theta_i = \overline{\theta}'_i, \theta_j \in \widehat{\Theta}_{j,l-1}\right\}$ (otherwise $\underline{\theta}'_i$ and $\overline{\theta}'_i$ would not be indifferent types). Also, $h_j\left(\underline{\theta}'_i\right) = \left\{\theta_j \in \Theta_j | v_j\left(\underline{\theta}'_i, \theta_j\right) = 0\right\} \in \left(\underline{\theta}'_j, \overline{\theta}'_j\right)$, i.e., $v_j(\theta) = 0$ goes through a vertical segment $\left\{\theta | \theta_i = \underline{\theta}'_i, \theta_j \in \widehat{\Theta}_{j,l}\right\}$ for the same reason. Hence, a mapping $t : \widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,l} \to \widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,l}$ defined by $t(x_i, x_j) = \left(h_i\left(\underline{\theta}'_j, x_j\right), h_j(\underline{\theta}'_i, x_i)\right)$ has at least one fixed point in the interior of $\widehat{\Theta}_{k,l}$. Let $x^* = (x^*_i, x^*_j) \in \widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,l}$ be such a fixed point. This is a $(2,2)$-rule defined on $\widehat{\Theta}_{i,k} \times \widehat{\Theta}_{j,k}$. Therefore, a (K, L)-rule with $K + L \geq 6$ is not PRP.

Note. When a $(2,2)$-rule exists in a posterior equilibrium of the first mechanism, it can always be made *posterior individual rational* in the following sense. For the $(2,2)$-rule constructed on $\widehat{\Theta}_{k,l}$ above, let the lower outcome be $\phi^-(< \phi^+)$ on the area below and the left to x^* and keep the original outcome ϕ^+ on the remaining area. Because agents agree on d_0 conditional on coordinated actions, ϕ^- is unanimously preferred to ϕ^+. On the other hand, there may be type-dependent disagreement when they do not coordinate their actions. However, agents are kept indifferent because ϕ^+ was not changed. Thus, the existence of $(2,2)$-rule in a posterior equilibrium is not merely a possibility of renegotiation, it actually is a case where improvement can be easily constructed. The same argument applies to any (K,K)-rules.

(ii) (If) If (a) holds, there is no $(1,2)$ and $(2,1)$-rule and by **Lemma 2** (i) there is no $(2,2)$-rule. By **Lemma 1 (i)**, there is no posterior implementable rule that is not constant. Because there is interim disagreement, for any $\phi^0 \in [0,1]$, unanimous improvement by a constant rule is impossible. If (b) holds, unanimous improvement by a constant rule is impossible because $\phi^0 = 1$. Also, there is no $(1,2)$ and $(2,1)$-rule. By the converse of **Lemma 2** (ii), there is no $(2,2)$-rule of high type. Because $(2,2)$-rule of low type does not exist by assumption, a constant rule $\phi^0 = 1$ is PRP. If (c) holds, a symmetric argument of (b) applies.

(Only if) If none of (a)-(c) holds, either (a') $(1,2)$ or $(2,1)$-rules exist, or (b') there is interim agreement on d_1 and ($\phi^0 < 1$ or $(2,2)$-rule of low type exists), or (c') there is interim agreement on d_0 and ($\phi^0 > 0$ or $(2,2)$-rule of high type exists). None of these cases satisfies the definition of PRP.

(iii) (If) If (a) holds, there is interim disagreement on $\{\theta_i < \theta_i^d\} \times \Theta_j$ and there is interim agreement on d_1 on $\{\theta_i > \theta_i^d\} \times \Theta_j$. By **Proposition (ii)** applied to each case, the requirement for PRP is satisfied. If (b) holds, a symmetric argument of (a) applies. If (c) holds, there is interim agreement on d_0 on $\{\theta_i < \theta_i^d\} \times \Theta_j$ and interim agreement on d_1 on $\{\theta_i > \theta_i^d\} \times \Theta_j$. There is no improvement by a constant rule because $\phi^- = 0$ on $\{\theta_i < \theta_i^d\} \times \Theta_j$ and $\phi^+ = 1$ on $\{\theta_i > \theta_i^d\} \times \Theta_j$. By the converse of **Lemma 2** (ii) applied to each case, there is no $(2,2)$-rule of low type on $\{\theta_i < \theta_i^d\} \times \Theta_j$ and there is no $(2,2)$-rule of high type on $\{\theta_i > \theta_i^d\} \times \Theta_j$. Thus, the last statement in (c) guarantees that there is no more information revelation in each case.

(Only if) If none of (a)-(c) holds, either (a') $h_j\left(\underline{\theta}_i, \theta_i^d\right) \in \left(\underline{\theta}_j, \overline{\theta}_j\right)$ or $h_j\left(\theta_i^d, \overline{\theta}_i\right) \in \left(\underline{\theta}_j, \overline{\theta}_j\right)$, (b') ($h_j\left(\underline{\theta}_i, \theta_i^d\right) = \underline{\theta}_j$ and $\phi^+ < 1$) or ($h_j\left(\theta_i^d, \overline{\theta}_i\right) = \overline{\theta}_j$ and $\phi^- > 0$), (c') $h_j\left(\underline{\theta}_i, \theta_i^d\right) = \overline{\theta}_j$, $h_j\left(\theta_i^d, \overline{\theta}_i\right) = \underline{\theta}_j$ and ($\phi^+ < 1$ or $\phi^- > 0$ or a $(2,2)$-rule exists either one of two partitions). None of these cases satisfies the definition of PRP.

(iv) (If) On $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$, $(2,1)$-rules cannot exist because $\theta_{i,1} = \min \widehat{\Theta}_{i,2}$ is an indifferent type given $\left\{\theta_j \in \widehat{\Theta}_{j,1}\right\}$. Similarly, on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,2}$, $(1,2)$-rules cannot exist because $\theta_{j,1} = \min \widehat{\Theta}_{j,2}$ is an indifferent type given $\left\{\theta_i \in \widehat{\Theta}_{i,1}\right\}$. On $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$, there is interim agreement on d_0. By the converse of **Lemma 2** (ii), there is no $(2,2)$-rule of high type on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$. Therefore, (a)-(c) is sufficient for no information revelation. Because there is interim agreement on d_0 on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$ and interim agreement on d_1 in the other partitions, $\phi^- = 0$ and $\phi^+ = 1$ imply there is no improvement by a constant rule.

(Only if) Necessity of (a)-(c) is obvious by the definition of PRP. If (d) is violated when (a)-(c) are satisfied, improvement by a constant rule is possible either on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$ (by lower outcome than $\phi^- > 0$) or on the other partitions (by higher outcome than $\phi^+ < 1$).

(v) (a) (If) First, note that in a $(3,2)$-rule $v_i(\theta) = 0$ crosses two vertical segments while $v_j(\theta) = 0$ crosses a horizontal segment. Because neither line goes through $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$ and $\widehat{\Theta}_{i,3} \times \widehat{\Theta}_{j,2}$, there cannot be any posterior implementable rule that is not constant. Second, on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,2}$, $(2,1)$-, $(1,2)$-, and $(2,2)$-rules of a high type do not exist (consider $\theta_{i,1}$ and $\theta_{j,1}$). Similarly, $(2,1)$-, $(1,2)$-, and $(2,2)$-rules of a low type do not exist on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$ (consider $\theta_{i,2}$ and $\theta_{j,1}$). Third, on $\widehat{\Theta}_{i,3} \times \widehat{\Theta}_{j,1}$, $(2,1)$- and $(2,2)$-rules of a high type do not exist (consider $\theta_{i,2}$). Similarly, $(2,1)$- and $(2,2)$-rules of a low type do not exist on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,2}$ (consider $\theta_{i,1}$). Hence, it suffices to show: if neither $(2,2)$-rule of a low type on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,2}$ nor $(2,2)$-rule of a high type on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$ exists, then there is neither $(1,2)$- nor $(2,2)$-rule of a low type on $\widehat{\Theta}_{i,3} \times \widehat{\Theta}_{j,1}$ and there is neither $(1,2)$- nor $(2,2)$-rule of a high type on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,2}$. Notice that if $v_j(\theta) = 0$ crosses the left vertical segment $\left\{\theta_{i,1}, \left(\theta_{j,1}, \overline{\theta}_j\right)\right\}$, then there exists a $(2,2)$-rule of a low type on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,2}$. Too see this, consider $h_i(\theta_{j,1}, x_j)$ defined for $x_j \in \left[\theta_{j,1}, \overline{\theta}_j\right]$, which connects $h_i\left(\theta_{j,1}, \overline{\theta}_j\right) = \theta_{i,1}$ and $h_i\left(\theta_{j,1}, \theta_{j,1}\right) = \left\{\theta_i \in \widehat{\Theta}_{i,2} | v_i\left(\theta_i, \theta_{j,1}\right) = 0\right\} \in$

$(\theta_{i,1}, \theta_{i,2})$. Also consider $h_j(\theta_{i,1}, x_i)$ defined for $x_i \in [\theta_{i,1}, \theta_{i,2}]$, which connects $h_j(\theta_{i,1}, \theta_{i,2}) = \theta_{j,1}$ and $h_j(\theta_{i,1}, \theta_{i,1}) = \left\{\theta_j \in \widehat{\Theta}_{j,2} | v_j(\theta_{i,1}, \theta_j) = 0\right\} \in \left(\theta_{j,1}, \overline{\theta}_j\right)$. They must cross at least once. Similarly, if $v_j(\theta) = 0$ crosses the right vertical segment $\left\{\theta_{i,2}, \left(\underline{\theta}_j, \theta_{j,1}\right)\right\}$, then there exists a $(2,2)$-rule of a high type on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$. Therefore, if neither $(2,2)$-rule of a low type on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,2}$ nor $(2,2)$-rule of a high type on $\widehat{\Theta}_{i,2} \times \widehat{\Theta}_{j,1}$ exists, then $v_j(\theta) = 0$ crosses none of the two vertical segments. This implies that $v_j(\theta) = 0$ crosses neither $\widehat{\Theta}_{i,3} \times \widehat{\Theta}_{j,1}$ nor $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,2}$. Hence, there can be neither $(1,2)$- nor $(2,2)$-rules on these type sets. Thus, (a)-(c) guarantee that there is interim agreement in each partition. $\phi^- = 0$ and $\phi^+ = 1$ imply there is no improvement by a constant rule.

(Only if) Necessity of non-existence of stated $(2,2)$-rules is obvious by the definition of PRP. That either $\phi^- > 0$ or $\phi^+ < 1$ clearly violates the definition of PRP because there is an improvement by a constant rule either on $\widehat{\Theta}_{i,3} \times \widehat{\Theta}_{j,2}$ (by higher outcome than ϕ^+) or on $\widehat{\Theta}_{i,1} \times \widehat{\Theta}_{j,1}$ (by lower outcome than ϕ^-).

(b) The condition says $v_j(\theta) = 0$ crosses none of the two vertical segments. The necessity of this condition was proved in (a). ■

Proof of Corollary.

(i) With symmetric agents, **Proposition (ii) (a)** cannot occur. By **Lemma 4 (i)(v)**, no information revelation is possible if and only if $\{\theta \in \Theta | v_i(\theta) = 0\} \cap \left(\underline{\theta}, \overline{\theta}\right)^2$ is empty. There are two cases to consider. If $v_i(\theta) \geq 0\ \forall \theta \in \left[\underline{\theta}, \overline{\theta}\right]$, then there is interim agreement on d_1. Hence, $\phi^0 = 1$ implies PRP and $\phi^0 < 1$ implies not PRP. If $v_i(\theta) \leq 0\ \forall \theta \in \left[\underline{\theta}, \overline{\theta}\right]$, then there is interim agreement on d_0. Hence, $\phi^0 = 0$ implies PRP and $\phi^0 > 0$ implies not PRP.

(ii) With symmetric agents, **Proposition (iii)(a)(b)** do not hold.

(iii) If $\theta_{i,1} = \theta_{j,1}$, two partitions $[\underline{\theta}, \theta_{i,1}] \times [\underline{\theta}, \theta_{j,1}]$ and $\left[\theta_{i,1}, \overline{\theta}\right] \times \left[\theta_{j,1}, \overline{\theta}\right]$ are both symmetric and agents' beliefs remain symmetric. If the $(2,2)$-rule is of low type, $v_i(\theta) = 0$ goes through the interior of $[\underline{\theta}, \theta_{i,1}] \times [\underline{\theta}, \theta_{j,1}]$, while if the rule is of high type, $v_i(\theta) = 0$ goes through the interior of $\left[\theta_{i,1}, \overline{\theta}\right] \times \left[\theta_{j,1}, \overline{\theta}\right]$. By **Lemma 4 (i)(v)**, a $(2,2)$-rule exists on either partition. ■

References

[1] Cramton, P.C. and T.R. Palfrey (1995): "Ratifiable Mechanisms: Learning from Disagreement," *Games and Economic Behavior*, 10, 255-283.

[2] Forges, F. (1994): "Posterior Efficiency," *Games and Economic Behavior*, 6, 238-261.

[3] Green, J. R. and J.J. Laffont (1987): "Posterior Implementability in a Two-person Decision Problem," *Econometrica*, 55, 1, 69-94.

[4] Holmstrom, B. and R. B. Myerson (1983): "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51, 6, 1799-1819.

[5] Jehiel, P., Meyer-ter-Vehn, M., Moldovanu, B., and W. Zame (2007): "Posterior Implementation vs Ex-post Implementation," *Economics Letters*, 97, 70-73.

[6] Lopomo, G. (2000): "Optimality and Robustness of the English Auction," *Games and Economic Behavior*, 36, 219-240.

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

[8] Vartiainen, H. (2013): "Auction Design without Commitment," *Journal of the European Economic Association*, 11, 2, 316-342.

[9] Watson, J. (1999): "Starting Small and Renegotiation," *Journal of Economic Theory*, 85, 52-90.