

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

Kei Kawakami[†]

April 4, 2015

Abstract

When two agents with private information use a mechanism to determine an outcome, what happens if they are free to revise their messages and cannot commit 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 conditions under 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 with at most five messages for two agents.

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

JEL classification code: D82, D83

*I would like to thank Georgy Artemov, Chris Edmond, and seminar participants at the 2013 North American Summer Meeting of Econometric Society, SAET 2014 (Tokyo), and anonymous referees and the associated editor for helpful comments on earlier drafts. 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) messages and (ii) a mechanism? 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 outcome to compare. In this environment, while there is no exogenous commitment to the first mechanism, the first mechanism has an informational 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 exogenous commitment to messages. We develop 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.¹ Green and Laffont (1987) relaxed commitment to messages 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 message profile. Green and Laffont interpret posterior implementable decision rules as representing a communication process without binding commitment to messages. Therefore, a posterior implementable decision rule must create “an incentive compatible information structure”, where it is common knowledge that no agent has an incentive to reconsider his message *available in the given mechanism*.

In this paper, we take this idea one step further by relaxing commitment to a mechanism. Suppose that 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

¹See Poitevin (2000) for a discussion of the Revelation Principle in this context.

any posterior implementable decision rule. Obviously the part (ii) is far from precise and many different notions of “robustness” can be considered. We take an extreme approach: 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 must create “a renegotiation-proof information structure”, where it is common knowledge that no mechanism can create a new incentive compatible information structure. Intuitively, a PRP decision rule leaves no possibility for future improvement based on information that can be revealed without commitment to messages.

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

We provide two motivations for our solution concept. First, the prospect of renegotiation affects the incentive compatibility of the original decision rule. 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.

Second, the concept 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 play a specified game). 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 services offered in the market will be of very limited variety and, to the contrary of the intension of the disclosure regulation, not much information will be revealed.

To our knowledge, our work is the first attempt to relax both commitments to messages and to 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.² 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. Section 3 presents examples. Section 4 presents the characterization of posterior implementation. Section 5 formally defines a PRP decision rule and characterizes it. Section 6 concludes. Section 7 contains proofs.

1.1 PRP and related solution concepts

We discuss key ingredients of our solution concept and compare it with related solution concepts. 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.

²See Lopomo (2000), Jehiel et al (2007), Vartiainen (2013) for recent studies of posterior implementation.

Table 1. Related solution concepts.

Info revelation	A. In a given mechanism	B. Comparison to alternatives
1. Individual optimality	<i>Posterior equilibrium</i> Green and Laffont (1987)	
2. Collective optimality	<i>Posterior efficiency</i> Forges (1994)	<i>Durability</i> Holmstrom and Myerson (1983) <i>Security</i> Cramton and Palfrey (1995)

Table 1 lists related solution concepts which capture *robustness of optimality to endogenous information revelation* in one way or another. Commitment to messages 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 for different cells of **Table 1**. Our solution concept addresses 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 a mechanism *and* by comparison to alternatives). 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 robustness. In terms of the timing, we allow agents to play two mechanisms and observe messages 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 messages. This timing assumption captures an informational advantage of the first mechanism: it can reveal information before the second mechanism is played. We ask the following question: Can we design a mechanism such that it only reveals information which “protects” itself? The answer depends on what “protects” means.

Consider the stage where the first mechanism was played, but no alternative has been proposed yet. Green and Laffont (1987) showed that in a posterior equilibrium the information structure must be coarse — the first mechanism cannot achieve the full information revelation. While this may leave much room for the second mechanism to reveal more information, it seems natural that agents lack commitment to messages in the second mechanism as well. Therefore, renegotiation should be represented by a posterior implementable rule. Then the problem boils down to the comparison of two posterior implementable decision rules. This comparison can be very complicated, because both the first and the second mechanisms can reveal some information, but not full information. Therefore, choosing one rule over the other based on some criteria (e.g. by unanimous voting) can potentially reveal new information. Accordingly, it seems inevitable that details of the renegotiation process matter. We address this problem by requiring that, conditional on the information revealed in the first mechanism, *only constant rules are posterior implementable*. In other words, the first mechanism is designed to ensure that no more information revelation is possible. The information revealed by the first mechanism “protects” itself because it creates common knowledge that any improvement by alternatives *must be* achieved without revealing further information. Once this can be guaranteed, it is straightforward to make the first decision rule renegotiation-proof – a posterior equilibrium must be Pareto efficient *with respect to its equilibrium information structure*.

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 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 messages reveal information, it is common knowledge that there can be no more information

revelation.³ Additionally, (ii) rules out 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*, then there must be information revelation. As long as this information revelation must occur in some game without commitment to messages, 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). In sum, we define PRP decision rules to have the following features: no more voluntary information revelation and Pareto efficiency with respect to the information revealed by the play of the first mechanism.

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 to 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 decision rules should be based on endogenously revealed information in the implementation process. The first difference is that in our solution concept a public message profile reveals information, while in posterior efficiency a public outcome reveals information. Hence, information revealed by messages but not by the outcome is ignored in posterior efficiency. Two concepts are also different in their timing assumptions. In posterior efficiency, comparison is made after the first outcome is known, but before the second mechanism is

³For example, this makes the first mechanism robust against the following possibility: agents play a second mechanism, and then *replay* the first mechanism.

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 the 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 *messages*), 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 equilibrium refinement of the unanimous voting game. 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 first mechanism itself does not reveal any information. Our solution concept may appear stronger than theirs by requiring that no information revelation occurs when comparing two outcomes. Again, however, because different endogenous information is considered in each approach, there is no obvious inclusion relationship.

2 Model

We study a collective decision problem in Green and Laffont (1987). We first explain pay-offs and information structure, where we add one technical assumption to their original environment. Second, we introduce a concept of posterior implementability.

Payoffs and information. Two agents, 1 and 2, have *type* $\theta = (\theta_1, \theta_2) \in \Theta_1 \times \Theta_2 = \Theta$. We assume that the type set is $\Theta_1 \times \Theta_2 = [\underline{\theta}_1, \bar{\theta}_1] \times [\underline{\theta}_2, \bar{\theta}_2] \subset \mathbb{R}^2$. The joint distribution function $F(\theta)$ and its density function $f(\theta)$ are common knowledge. Agents 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 1 is $u_1(d, \theta) = v_1(\theta)\mathbf{1}\{d = d_1\}$, and the payoff of agent 2 is $u_2(d, \theta) = v_2(\theta)\mathbf{1}\{d = d_1\}$.⁴ We call $v_1(\theta)$ and $v_2(\theta)$ *payoff functions*. A *decision rule* $\phi : \Theta \rightarrow [0, 1]$ associates any type θ with an *outcome* $\phi(\theta)$, which is the probability of d_1 . We simply use a *rule* instead of a decision rule throughout the paper. The following two assumptions are from Green and Laffont (1987).

Assumption 1: (a) $v_i(\theta)$ is continuous and strictly increasing in both arguments, $i = 1, 2$.

(b) The set $\{\theta \in \Theta | v_1(\theta) = v_2(\theta) = 0\}$ 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 , $i = 1, 2$, $j \neq i$.

(b) For any subinterval $\hat{\Theta}_j$ of Θ_j , the conditional distributions $F_i(\theta_j | \theta_i, \theta_j \in \hat{\Theta}_j)$ satisfy $\theta_i < \theta'_i \Rightarrow F_i(x | \theta_i, \theta_j \in \hat{\Theta}_j) \geq F_i(x | \theta'_i, \theta_j \in \hat{\Theta}_j) \quad \forall x \in \hat{\Theta}_j$, for $i = 1, 2$, $j \neq i$.

Note that $\{\theta \in \Theta | v_1(\theta) = 0\}$ is a subset of the type set Θ for which agent 1 is indifferent between d_0 and d_1 . Due to Assumption 1, $\hat{\theta}_1(\theta_2) \equiv \{\theta_1 \in \Theta_1 | v_1(\theta) = 0\}$ is a strictly decreasing continuous function. Throughout the paper, we measure θ_1 in the horizontal direction and θ_2 in the vertical direction so that “a vertical segment” means $\{\theta \in \Theta | \theta_1 = a, \theta_2 \in [b, c]\}$, while “a horizontal segment” means $\{\theta \in \Theta | \theta_1 \in [a, b], \theta_2 = c\}$.

⁴Throughout the paper, we use $\mathbf{1}\{X\}$ to denote an indicator function that returns one if and only if the statement X is true.

Figure 1. Indifference curves.

Figure 1 illustrates Assumption 1: two indifference curves $v_1(\theta) = 0$ and $v_2(\theta) = 0$ are strictly decreasing and intersect at most finite number of times. Both agents prefer d_1 to d_0 in the area above the two curves, while both prefer d_0 to d_1 in the area below the two curves. Hence two agents agree in these areas under complete information. In the area labeled as $D1$, agent 1 prefers d_1 to d_0 while agent 2 prefers d_0 to d_1 . The opposite is true in $D2$. Hence two agents disagree in $D1$ and $D2$ under complete information.⁵ An efficient 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.

Consider the expected payoff from d_1 for agent 1 with type θ_1 believing that agent 2's type is in a subinterval $\hat{\Theta}_2$ of Θ_2 :

$$V_1(\theta_1; \hat{\Theta}_2) \equiv \begin{cases} \int_{\Theta_2} v_1(\theta_1, y) dF_1(\theta_2 | \theta_1, y \in \hat{\Theta}_2) & \text{if } \hat{\Theta}_2 \text{ is an interval,} \\ v_1(\theta_1, \theta_2) & \text{if } \hat{\Theta}_2 = \{\theta_2\}. \end{cases} \quad (1)$$

We define $V_2(\theta_2; \hat{\Theta}_1)$ by switching the role of 1 and 2 in (1). Assumption 1 and Assumption 2(a) make (1) well-defined for all θ_1 and any subinterval $\hat{\Theta}_2$ of Θ_2 . Assumption 2(b)

⁵More 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 excluding the intersections, and both agree in the remaining area.

introduces a positive correlation of types, which makes (1) increasing in θ_1 . Because of the monotonicity, the set $\{\theta_1 \in \Theta_1 | V_1(\theta_1; \widehat{\Theta}_2) = 0\}$ has only one element if it is not empty. For any subintervals $[x_1, y_1] \subseteq \Theta_1$ and $[x_2, y_2] \subseteq \Theta_2$, we define two functions:

$$\begin{aligned} h_1(x_2, y_2) &\equiv \begin{cases} \{\theta_1 \in \Theta_1 | V_1(\theta_1; [x_2, y_2]) = 0\} & \text{if the set is not empty,} \\ \underline{\theta}_1 & \text{if } V_1(\theta_1; [x_2, y_2]) > 0 \ \forall \theta_1 \in \Theta_1 \\ \bar{\theta}_1 & \text{if } V_1(\theta_1; [x_2, y_2]) < 0 \ \forall \theta_1 \in \Theta_1. \end{cases} \\ h_2(x_1, y_1) &\equiv \begin{cases} \{\theta_2 \in \Theta_2 | V_2(\theta_2; [x_1, y_1]) = 0\} & \text{if the set is not empty,} \\ \underline{\theta}_2 & \text{if } V_2(\theta_2; [x_1, y_1]) > 0 \ \forall \theta_2 \in \Theta_2 \\ \bar{\theta}_2 & \text{if } V_2(\theta_2; [x_1, y_1]) < 0 \ \forall \theta_2 \in \Theta_2. \end{cases} \end{aligned} \quad (2)$$

While $V_1(\theta_1; \widehat{\Theta}_2)$ is well-defined and increasing in θ_1 , it may be discontinuous in θ_1 .⁶ To make sure that h_1 and h_2 are well-defined and continuous, we add the following assumption.

Assumption 3: For any subinterval $\widehat{\Theta}_j$ of Θ_j , $V_i(\theta_i; \widehat{\Theta}_j)$ is continuous in θ_i , $i = 1, 2$, $j \neq i$.

The function $h_1 : \Theta_2 \times \Theta_2 \rightarrow \Theta_1$ determines an indifference type of agent 1 who believes that agent 2's type lies in the set $[x_2, y_2] \subseteq \Theta_2$ specified by two inputs of the function. Compared with a function $\widehat{\theta}_1(\theta_2) \equiv \{\theta_1 \in \Theta_1 | v_1(\theta) = 0\}$, which determines an indifference type of agent 1 based on *the ex post belief*, h_1 can be interpreted as a generalization of $\widehat{\theta}_1$ where agent 1's belief can be coarser than the ex post belief.

Remark on Assumption 3. When types are independent, Assumption 3 is satisfied. Intuitively, to satisfy Assumption 3, conditional distributions $F_1(\theta_2 | \theta_1, \theta_2 \in \widehat{\Theta}_2)$ must shift continuously as θ_1 increases. In the next section, we present a family of distributions which satisfies Assumptions 2 and 3. With these assumptions, h_1 and h_2 defined in (2) are continuous, and whenever they return the values other than $\underline{\theta}_1$, $\bar{\theta}_1$, $\underline{\theta}_2$, $\bar{\theta}_2$, they strictly de-

⁶This may occur when an increase in θ_1 results in the discontinuous shift in the conditional density $f_1(\cdot | \theta_1)$. This can lead to the possibility that $V_1(\theta_1; [x_2, y_2]) < 0 \ \forall \theta_1 \in [\underline{\theta}_1, \theta^*]$ and $V_1(\theta_1; [x_2, y_2]) > 0 \ \forall \theta_1 \in (\theta^*, \bar{\theta}_1]$. I appreciate an anonymous referee for this point.

crease in both arguments. Green and Laffont (1987) did not require Assumption 3, because they only characterized posterior implementable rules without establishing their existence.⁷ Assumption 3 is crucial for our analysis because the existence/non-existence of posterior implementable rules is the key for posterior renegotiation-proofness.

Posterior implementation. A *mechanism* (M, g) is a measurable set $M = M_1 \times M_2$ and a measurable function $g : M \rightarrow [0, 1]$, where $g(m)$ is a probability of d_1 when messages $m = (m_1, m_2) \in M$ are chosen by agents. A *strategy* of agent 1 is a collection of conditional distributions $s_1(m_1|\theta_1)$, $\theta_1 \in \Theta_1$. A pair of strategies $s = (s_1, s_2)$ in the mechanism (M, g) results in the rule $\phi(\theta) = \int_M g(m) ds_1(m_1|\theta_1) ds_2(m_2|\theta_2)$. Let $\mu^s(m, \theta)$ be the joint distribution over $M \times \Theta$ generated in a Bayes-Nash equilibrium s .⁸ Let $\mu_1^s(m_2|\theta_1)$ be the marginal distribution of m_2 given θ_1 . For every θ_1 and $\mu_1^s(m_2|\theta_1)$ -almost every m_2 , define $F_1(\theta_2|\theta_1, m_2)$ to be the conditional distribution that agent 1 with type θ_1 would hold about θ_2 given m_2 . Define $\mu_2^s(m_1|\theta_2)$ and $F_2(\theta_1|\theta_2, m_1)$ by switching the role of 1 and 2.

Definition 1: A Bayes-Nash equilibrium s is a posterior equilibrium if

$$\begin{aligned} m_1 &\in \arg \max_{m'_1 \in M_1} g(m'_1, m_2) \int_{\Theta_2} v_1(\theta_1, \theta_2) dF_1(\theta_2|\theta_1, m_2), \\ m_2 &\in \arg \max_{m'_2 \in M_2} g(m_1, m'_2) \int_{\Theta_1} v_2(\theta_1, \theta_2) dF_2(\theta_1|\theta_2, m_1), \end{aligned} \tag{3}$$

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

Definition 2: A rule ϕ is posterior implementable if there is a mechanism (M, g) with a posterior equilibrium s which results in ϕ .

The optimality condition (3) shows that a message m_1 affects an equilibrium outcome through two channels. The first channel is directly through the function g . The second

⁷Their main result is stated as a necessary condition for posterior implementable rules. However, on page 88 of Green and Laffont (1987), they applied Brouwer's fixed-point theorem using functions h_1 and h_2 defined here. It appears they assumed the continuity of these functions in some part of their analysis.

⁸The formal definition of the Bayes-Nash equilibrium was presented in Green and Laffont (1987) and the earlier version of this paper. Since it does not add much to our analysis, we omit it here.

channel is indirectly through $F_2(\theta_1|\theta_2, m_1)$, i.e., by affecting the belief of agent 2 and hence m_2 . Therefore, for a posterior equilibrium to exist, not only g but also the message space $M_1 \times M_2$ must be carefully designed. Identifying these necessary conditions was the contribution of Green and Laffont (1987), which we review in Section 4. In the next section, we present concrete examples of posterior implementable rules and PRP rules.

3 Examples

This section studies examples with the following linear payoff functions:

$$\begin{aligned} v_1(\theta) &= \theta_1 + a_1 (\theta_2 - c_1), & a_1 > 0, c_1 > 0, \\ v_2(\theta) &= \theta_2 + a_2 (\theta_1 - c_2), & a_2 > 0, c_2 > 0, \end{aligned} \tag{4}$$

and the following family of conditional density functions defined on $\Theta = [0, 1]^2$:

$$\begin{aligned} \text{For } \theta_1 = 0, & \quad f_1(\theta_2|\theta_1) = \varepsilon + 2(1 - \varepsilon)(1 - \theta_2). \\ \text{For } \theta_1 \in (0, 1), & \quad f_1(\theta_2|\theta_1) = \begin{cases} \varepsilon + 2(1 - \varepsilon)\frac{\theta_2}{\theta_1} & \text{if } \theta_2 \leq \theta_1 \\ \varepsilon + 2(1 - \varepsilon)\frac{1 - \theta_2}{1 - \theta_1} & \text{if } \theta_2 > \theta_1 \end{cases}. \\ \text{For } \theta_1 = 1, & \quad f_1(\theta_2|\theta_1) = \varepsilon + 2(1 - \varepsilon)\theta_2, \end{aligned}$$

where $\varepsilon \in (0, 1]$.⁹ With $\varepsilon = 1$, $f_1(\theta_2|\theta_1)$ is a uniform distribution on $[0, 1]$ independent of θ_1 . With $\varepsilon < 1$, $f_1(\theta_2|\theta_1)$ is linear in θ_2 and has a single peak at $\theta_2 = \theta_1$, which continuously shifts to the right as θ_1 increases. As ε decreases, the probability mass that moves with θ_1 increases, capturing the idea that agents believe they are more “alike”. This family of distributions satisfies Assumption 2, because for any subinterval $\hat{\Theta}_2$ of Θ_2 , $f_1(\theta_2|\theta_1, \theta_2 \in \hat{\Theta}_2)$ has a single peak which continuously shifts to the right as θ_1 increases. Assumption 3 can be directly verified as we show below.

⁹The assumption $\Theta_1 = \Theta_2 = [0, 1]$ can be relaxed with the burden of added notations.

3.1 Example 1: Symmetric agents

We first illustrate posterior implementable rules which invite renegotiation. Assume $a_1 = a_2 = a > 0$ and $c_1 = c_2 = 1$ in (4). This symmetric linear preference implies that the two indifference curves intersect only once at $(\frac{a}{1+a}, \frac{a}{1+a}) \in \Theta$. We can explicitly compute:

$$V_1(\theta_1; [0, y]) = \begin{cases} \int_0^{\theta_1} \{\theta_1 + a(\theta_2 - 1)\} \left\{ \varepsilon + 2(1 - \varepsilon) \frac{\theta_2}{\theta_1} \right\} d\theta_2 & \text{if } y > \theta_1, \\ + \int_{\theta_1}^y \{\theta_1 + a(\theta_2 - 1)\} \left\{ \varepsilon + 2(1 - \varepsilon) \frac{1 - \theta_2}{1 - \theta_1} \right\} d\theta_2 & \\ \int_0^y \{\theta_1 + a(\theta_2 - 1)\} \left\{ \varepsilon + 2(1 - \varepsilon) \frac{\theta_2}{\theta_1} \right\} d\theta_2 & \text{if } y \leq \theta_1. \end{cases}$$

First, with $y = 1$, $V_1(\theta_1; [0, 1]) = \frac{1}{6} \{(6 + 2a(1 - \varepsilon))\theta_1 - a(4 - \varepsilon)\}$. Since $h_1(0, 1) = 1$ if $V_1(\theta_1; [0, 1]) < 0$ for all $\theta_1 \in [0, 1]$,

$$h_1(0, 1) \equiv \theta^d(a, \varepsilon) = \begin{cases} \frac{4 - \varepsilon}{\frac{6}{a} + 2 - 2\varepsilon} & \text{if } a < \frac{6}{2 + \varepsilon} \equiv \bar{a}(\varepsilon), \\ 1 & \text{else.} \end{cases}$$

This is an indifferent type of agent 1 given his interim belief that $\theta_2 \in [0, 1]$.

Next, computing $V_1(\theta_1; [0, y])$ for $y \leq \theta_1$,

$$h_1(0, y) = \left\{ \theta_1 \in [0, 1] \mid \varepsilon \theta_1^2 - \left\{ \varepsilon a \left(1 - \frac{y}{2}\right) - (1 - \varepsilon)y \right\} \theta_1 - (1 - \varepsilon)a \left(1 - \frac{2y}{3}\right) y = 0 \right\}.$$

The equation in the right hand side has a unique positive solution $\theta_1(y)$ which is decreasing in y . This is an indifferent type of agent 1 given the belief that agent 2's type is in the set $[0, y]$. By imposing $\theta_1 = y$ in the equation above, the unique positive solution to $h_1(0, y) = y$ is $\frac{6}{\frac{6}{a} + 4 - \varepsilon}$.¹⁰ We denote

$$\theta^L(a, \varepsilon) \equiv \begin{cases} \frac{6}{\frac{6}{a} + 4 - \varepsilon} & \text{if } a < \bar{a}(\varepsilon), \\ 1 & \text{else.} \end{cases}.$$

Note that $\theta^d(a, \varepsilon) \leq \theta^L(a, \varepsilon)$ with equality if and only if $a \geq \bar{a}(\varepsilon)$.

¹⁰With $\theta_1 = y$, the equation becomes $\varepsilon y - \left\{ \varepsilon a \left(1 - \frac{y}{2}\right) - (1 - \varepsilon)y \right\} - (1 - \varepsilon)a \left(1 - \frac{2y}{3}\right) = 0$.

Finally, computing $V_1(\theta_1; [x, 1]) = 0$ yields a decreasing function $h_1(x, 1)$. This is an indifferent type of agent 1 given the belief that agent 2's type is in the set $[x, 1]$. By solving for a unique solution of $h_1(x, 1) = x$, we obtain

$$\theta^H(a, \varepsilon) \equiv \frac{4 - \varepsilon}{\frac{6}{a} + 4 - \varepsilon} < \theta^L(a, \varepsilon).$$

Clearly, $\theta^d(a, \varepsilon)$, $\theta^L(a, \varepsilon)$, $\theta^H(a, \varepsilon)$ all increase in a . If we assume $a = \frac{2}{3}$ and $\varepsilon = 1$, we have $(\theta^H, \theta^d, \theta^L) = (\frac{1}{4}, \frac{1}{3}, \frac{1}{2})$. **Figure 2** below illustrates this.

Note. The left panel shows that θ^H is the indifference type believing that the other agent's type is in the set $[\theta^H, 1]$. The right panel shows that θ^L is the indifference type believing that the other agent's type is in the set $[0, \theta^L]$.

Figure 2. Determination of θ^d , θ^L , θ^H .

The left panel illustrates the derivation of θ^H . The decreasing function $h_1(x_2, 1)$ connects two points $(\theta^d, 0)$ and $(0, 1)$ as x_2 increases from zero to one.¹¹ Because of the symmetric linear structure, $h_1(x_2, 1)$ and $h_2(x_1, 1)$ have a unique intersection on the 45 degree line.¹² By construction, agent 1 with type θ^H believing that agent 2's type is in the set $[\theta^H, 1]$ is indifferent for any outcomes. Analogously, agent 2 with type θ^H believing that agent 1's

¹¹This is because $h_1(0, 1)$ defines θ^d and $h_1(1, 1) = \{\theta_1 \in \Theta_1 | v_1(\theta_1, 1) = 0\} = 0$.

¹²In general, (θ_1^H, θ_2^H) is characterized by two equations $\theta_1^H = h_1(\theta_2^H, 1)$ and $\theta_2^H = h_2(\theta_1^H, 1)$. In this symmetric example, $h_1(x_2, 1)$ and $h_2(x_1, 1)$ are identical and we can find $\theta_1^H = \theta_2^H = \theta^H$ by solving the single equation $x = h_1(x, 1)$.

type is in the set $[\theta^H, 1]$ is indifferent. Types greater than θ^H strictly prefer higher outcomes to lower outcomes, while types smaller than θ^H prefer the opposite.

Consider a mechanism with two messages $\{L, H\}$, which yields a high outcome if both agents choose H , and a low outcome otherwise. This mechanism has a posterior equilibrium: agents with type in the set $[0, \theta^H]$ choose L , while agents with type in the set $[\theta^H, 1]$ choose H . Given the equilibrium observation of H chosen by the other agent, agent with type in the set $[0, \theta^H]$ prefers a low outcome, while agent with type in the set $[\theta^H, 1]$ prefers a high outcome. Hence, there is no incentive to change messages. Given the equilibrium observation of L chosen by the other agent, the indifferent type *after this observation* lies between θ^d and a .¹³ Therefore, for agent choosing H , there exists a positive measure of types which strictly prefer a high outcome to a low outcome after observing a message L by the other agent. However, since changing the message does not affect the outcome, posterior incentive compatibility is satisfied. Therefore, the rule which assigns a high outcome in the area $[\theta^H, 1] \times [\theta^H, 1]$ and a low outcome elsewhere is posterior implementable. Similarly, the rule which assigns a low outcome in the area $[0, \theta^L] \times [0, \theta^L]$ and a high outcome elsewhere is posterior implementable.

Four types of posterior implementable rules.

The way $\theta^d(a, \varepsilon)$, $\theta^L(a, \varepsilon)$, $\theta^H(a, \varepsilon)$ are constructed shows that there are at least four types of posterior equilibria when $a < \bar{a}(\varepsilon)$. Let $\phi^- < \phi^+$ be two different levels of outcomes.

$$(1-1) \ M = \{L, H\} \times \{\emptyset\}, \ g(L, \emptyset) = \phi^-, \ g(H, \emptyset) = \phi^+.$$

In a posterior equilibrium, agent 1 with type $\theta_1 < \theta^d(a, \varepsilon)$ chooses L while agent 1

¹³Recall that $\theta^d = h_1(0, 1)$ is the indifference type believing that the other agent's type is in the set $[0, 1]$. For a given type, changing this belief to $[0, \theta^H]$ *decreases* the expected payoff. Hence, the indifference type associated with $[0, \theta^H]$ must be larger than θ^d , i.e., $h_1(0, \theta^H) > h_1(0, 1)$. Similarly, note that $a = h_1(0, 0)$ is the indifferent type believing that the other agent's type is 0. Changing this belief to $[0, \theta^H]$ *increases* the expected payoff. Hence, the indifference type associated with $[0, \theta^H]$ must be smaller than a , i.e., $h_1(0, \theta^H) < h_1(0, 0)$.

with type $\theta_1 > \theta^d(a, \varepsilon)$ chooses H .¹⁴

(1-2) $M = \{\emptyset\} \times \{L, H\}$, $g(\emptyset, L) = \phi^-$, $g(\emptyset, H) = \phi^+$.

In a posterior equilibrium, agent 2 with type $\theta_2 < \theta^d(a, \varepsilon)$ chooses L while agent 2 with type $\theta_2 > \theta^d(a, \varepsilon)$ chooses H .

(2-L) $M = \{L, H\}^2$, $g(L, L) = \phi^-$, $g(L, H) = g(H, L) = g(H, H) = \phi^+$.

In a posterior equilibrium, agents with type smaller than $\theta^L(a, \varepsilon)$ choose L , while agents with type larger than $\theta^L(a, \varepsilon)$ choose H .

(2-H) $M = \{L, H\}^2$, $g(H, H) = \phi^+$, $g(L, H) = g(H, L) = g(L, L) = \phi^-$.

In a posterior equilibrium, agents with type smaller than $\theta^H(a, \varepsilon)$ choose L , while agents with type larger than $\theta^H(a, \varepsilon)$ choose H .

In the cases (1-1) and (1-2), one agent chooses the outcome ignoring the other agent's type, and they form mirror images to each other with respect to a role of two agents. In the cases (2-L) and (2-H), one outcome requires the same message from both agents, while the other outcome occurs for all other scenarios. These cases form mirror images to each other with respect to an outcome that requires a coordination. In the subsequent analysis, these four posterior implementable rules play a key role. In particular, we show that these rules provide necessary conditions for information revelation: if there is *any* posterior implementable rule which is not constant, we can find one of these rules.

Posterior renegotiation. Next, we explain why the rules in **Figure 2** are not renegotiation-proof. Informally, a rule is not PRP if, for some observation of equilibrium messages, there is a posterior implementable rule with information revelation.

¹⁴Because agent 1 with type $\theta^d(a, \varepsilon)$ is indifferent, he can randomly choose either message.

Note. θ^H is the indifferent type believing that the other agent's type is in the set $[\theta^H, 1]$.

Figure 3. Posterior renegotiation.

In **Figure 3**, the observation of $m = (H, H)$ in a posterior equilibrium creates a *posterior type set* $[\theta^H, 1] \times [\theta^H, 1]$. Because $h_1(\theta^H, \theta^H) = b$ and $h_1(\theta^H, 1) = \theta^H$, $h_1(\theta^H, y_2)$ connects two points (b, θ^H) and $(\theta^H, 1)$ as y_2 increases from θ^H to one. Due to symmetry, $h_1(\theta^H, y_2)$ and $h_2(\theta^H, y_1)$ has a unique intersection on the 45 degree line. By construction, the case (2-L) applies on the set $[\theta^H, 1] \times [\theta^H, 1]$. Moreover, since the first mechanism assigned the high outcome ϕ^+ , the new mechanism can achieve Pareto improvement by offering a lower outcome only for $m = (L, L)$. Thus, the first rule is not renegotiation proof.¹⁵

3.2 Example 2: Asymmetric agents

We go back to a general linear payoff (4) and assume $a_1 = a > 4$, $a_2 = 2$, $c_1 = \frac{1}{2}$, $c_2 = \frac{3}{4}$, $\varepsilon = 1$. Note that agents are asymmetric. We show that the case (2-L) applies for all

¹⁵In this example, it can also be shown that the case (2-H) applies in the posterior type set $[\theta^H, 1] \times [0, \theta^H]$ after observing $m = (H, L)$, and also in the posterior type set $[0, \theta^H] \times [\theta^H, 1]$ after observing $m = (L, H)$.

$a \in (4, \infty)$ but it is PRP only if $a \geq 6$. First, compute

$$\begin{aligned} V_1(\theta_1; [0, y_2]) &= \int_0^{y_2} \left\{ \theta_1 + a \left(\theta_2 - \frac{1}{2} \right) \right\} d\theta_2 = y_2 \left[\theta_1 - \frac{a}{2} + \frac{a}{2} y_2 \right]. \\ V_2(\theta_2; [0, y_1]) &= \int_0^{y_1} \left\{ \theta_2 + 2 \left(\theta_1 - \frac{3}{4} \right) \right\} d\theta_1 = y_1 \left[\theta_2 - \frac{3}{2} + y_1 \right]. \end{aligned}$$

$V_1(\theta_1; [0, y_2]) = 0$ and $V_2(\theta_2; [0, y_1]) = 0$ define two decreasing functions $h_1(0, y_2) = \frac{a}{2} - \frac{a}{2}y_2$ and $h_2(0, y_1) = \frac{3}{2} - y_1$. Note that $h_1(0, y_2)$ connects two points $(0, 1)$ and $(1, 1 - \frac{2}{a})$, while $h_2(0, y_1)$ connects two points $(\frac{1}{2}, 1)$ and $(1, \frac{1}{2})$. Given $a > 4$, two equations $h_1(0, y_2) = y_1$ and $h_2(0, y_1) = y_2$ have a unique solution $(y_1, y_2) = \theta^L(a) \equiv \left(\frac{a}{2(a-2)}, \frac{a-3}{a-2} \right)$ in the interior of $[0, 1]^2$. As a moves from 4 to infinity, $h_1(0, y_2)$ rotates upward, moving $\theta^L(a)$ from $(1, \frac{1}{2})$ to $(\frac{1}{2}, 1)$ along $h_2(0, y_1)$, i.e., along the line $\theta_1 = \frac{3}{2} - \theta_2$. Thus, for any $a \in (4, \infty)$, the case (2-L) applies. **Figure 4** illustrates three cases, $a = 6$, $4 < a < 6$, and $a > 6$.

Note. The left panel is for $a = 6$. The middle panel is for $a \in (4, 6)$. The right panel is for $a > 6$. In the middle and right panels, thin lines are for $a = 6$ and arrows indicate the movement induced by a change in a from 6.

Figure 4. Posterior implementable rules and renegotiation.

To see why $a \geq 6$ is necessary for posterior renegotiation-proofness, consider $a = 6$ where $\theta^L = (\frac{3}{4}, \frac{3}{4})$ (the left panel). In the posterior type set $[0, \frac{3}{4}] \times [0, \frac{3}{4}]$, by computing

$$\int_{x_2}^{\frac{3}{4}} \{\theta_1 - 3 + 6\theta_2\} d\theta_2 = \left(\frac{3}{4} - x_2\right) \left(\theta_1 - \frac{3}{4} + 3x_2\right),$$

$$\int_{x_1}^{\frac{3}{4}} \left\{\theta_2 - \frac{3}{2} + 2\theta_1\right\} d\theta_1 = \left(\frac{3}{4} - x_1\right) \left(\theta_2 - \frac{3}{4} + x_1\right),$$

we obtain

$$h_1\left(x_2, \frac{3}{4}\right) = \frac{3}{4} - 3x_2 \text{ and } h_2\left(x_1, \frac{3}{4}\right) = \frac{3}{4} - x_1.$$

These are two double-dashed lines. Notice that $h_2(x_1, \frac{3}{4})$ is above $h_1(x_2, \frac{3}{4})$ except at the point $(\frac{3}{4}, 0)$ where they intersect. It can be easily verified that none of the four cases applies in the posterior type set $[0, \frac{3}{4}] \times [0, \frac{3}{4}]$.

Now consider the middle panel where a is slightly reduced from 6. Because $v_1(\theta) = \theta_1 + a(\theta_2 - \frac{1}{2}) = 0$ connects $(0, \frac{1}{2})$ and $(1, \frac{1}{2} - \frac{1}{a})$, it rotates down around $(0, \frac{1}{2})$. Therefore, $h_1(0, y_2)$ also rotates down (see the red dashed line), moving $\theta^L(a)$ down along the line $\theta_1 = \frac{3}{2} - \theta_2$. Next, $h_1(x_2, \theta_2^L(a))$ defined on the posterior type set $[0, \theta_1^L(a)] \times [0, \theta_2^L(a)]$ connects two points $(\theta_1^L(a), 0)$ and $(0, x^*)$, where x^* solves $h_1(x^*, \theta_2^L(a)) = 0$ (see the red double-dashed line). Since x^* and $\theta_2^L(a)$ must move in the opposite directions, x^* increases in response to a drop in a . Because both $\theta_1^L(a)$ and x^* increase, $h_1(x_2, \theta_2^L(a))$ moves up. The opposite comparative statics applies to $h_2(x_1, \theta_1^L(a))$ and it moves down.¹⁶ Therefore, for $a < 6$, $h_1(x_2, \theta_2^L(a))$ and $h_2(x_1, \theta_1^L(a))$ have an intersection in the posterior type set $[0, \theta_1^L(a)] \times [0, \theta_2^L(a)]$, and the case (2-H) applies. Therefore, the original posterior implementable rule is not renegotiation-proof.

Finally, consider the right panel where a is slightly raised from 6. This shifts $\theta^L(a)$ up along the line $\theta_1 = \frac{3}{2} - \theta_2$. Since $h_1(\cdot, \theta_2^L(a))$ moves down and $h_2(\cdot, \theta_1^L(a))$ moves up, they

¹⁶See the green double-dashed line. $h_2(x_1, \theta_1^L(a))$ connects two points $(0, \theta_2^L(a))$ and $(y^*, 0)$, where y^* solves $h_2(y^*, \theta_1^L(a)) = 0$. Because both θ_2^L and y^* decrease, $h_2(x_1, \theta_1^L(a))$ moves down.

move away from each other. Hence, they do not intersect in the set $[0, \theta_1^L(a)] \times [0, \theta_2^L(a)]$ and the case (2-H) does not apply. In fact, it can be shown that *none of the four cases applies in every posterior type set* if $a \geq 6$. This is an example of PRP rules.

4 Structure of posterior implementation

In this section, we go back to a general case and establish useful properties of posterior implementable rules. We first review a necessary condition for posterior implementable rules established in Green and Laffont (1987). Second, we present new characterization results which uncover the importance of (2, 1) and (2, 2)-rules for posterior implementation.

Characterization by Green and Laffont. Let H be the set of decreasing step functions ξ which partition $\Theta_1 \times \Theta_2$ in two parts with the following properties:

$$\begin{aligned} \text{any vertical segment } (\theta_1, [a, b]) \text{ of } \xi \text{ satisfies } \theta_1 &= h_1(a, b), \\ \text{any horizontal segment } ([c, d], \theta_2) \text{ of } \xi \text{ satisfies } \theta_2 &= h_2(c, d). \end{aligned} \tag{5}$$

Theorem (Green and Laffont, 1987)

Any posterior implementable rule ϕ is such that, for some $\xi \in H$, $\phi(\theta) = \phi^- \mathbf{1}\{\theta \text{ is below } \xi\} + \phi^+ \mathbf{1}\{\theta \text{ is above } \xi\}$ 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 the values of ϕ over the set of measure zero.¹⁷ **Figure 5** shows an example.

¹⁷The values of ϕ on ξ can take many values due to potential randomization by indifferent agents. In our model, the set of indifferent types have measure zero, so we can ignore the impact of their particular randomization choice on expected payoffs and hence on equilibrium strategies for other types.

Figure 5. A posterior implementable rule with three messages for each agent.

In **Figure 5**, each agent has three messages, each of which corresponds to a subinterval of types who uses that message in a posterior equilibrium. Each line segment in the step function has an associated indifferent type, characterized by one of the conditions in (5). For example, $\theta_{1,1} = h_1(\theta_{2,1}, \theta_{2,2}) = \{\theta_1 \in \Theta_1 | V_1(\theta_1; [\theta_{2,1}, \theta_{2,2}]) = 0\}$ so that agent 1 with type $\theta_{1,1}$ is indifferent between $m_{1,1}$ and $m_{1,2}$ after observing $m_{2,2}$.

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 $(\widehat{M}, \widehat{g})$ in terms of the resulting rule and the equilibrium information structure. A message space $\widehat{M}$ is a set of subsets of Θ . A message $\widehat{\Theta}_1 \in \widehat{M}_1$ by agent 1 is interpreted as “my type is in $\widehat{\Theta}_1$ ”, and it is truthful if his true type is in $\widehat{\Theta}_1$.¹⁸ Moreover, we can assume without loss of generality that each message is a closed interval and any two messages share at most one point.

For the rest of the paper, we use the following convention. Whenever a mechanism has multiple messages for agent 1, we order them such that $\max \widehat{\Theta}_{1,k} = \min \widehat{\Theta}_{1,k+1}$ for k th and $k+1$ th messages, and denote their boundary by $\theta_{1,k}$. If a mechanism has K messages available for agent 1 and L messages for agent 2, we call it a (K, L) -mechanism, and the associated rule shall be called a (K, L) -rule. Due to **Theorem** above, for any posterior implementable (K, L) -rules, K and L differ at most by one. We denote by $\Theta_\phi = \left\{ \widehat{\Theta}_{1,k} \times \widehat{\Theta}_{2,l} \right\}_{k=1, \dots, K, l=1, \dots, L}$

¹⁸For example, if $\Theta_1 \in \widehat{M}_1$, “my type is in Θ_1 ” is truthful although it reveals no information.

a partition of Θ created by a posterior implementable (K, L) -rule ϕ , and call each set $\hat{\Theta}_{k,l} \equiv \hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l}$ a *posterior type set*. While the set of posterior implementable rules may look restrictive, K and L can be 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.

4.1 New characterization

We provide new characterization of posterior implementable 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 rules defined in the next section. First, the following definitions will be useful for our purpose.

Definition 3: A constant mechanism is $M_1 = \{\Theta_1\}$, $M_2 = \{\Theta_2\}$, and $g(m) = \phi^0 \in [0, 1]$.

Definition 4: A dictatorial mechanism for agent 1 is $M_1 = \{\hat{\Theta}_{1,1}, \hat{\Theta}_{1,2}\}$, $M_2 = \{\Theta_2\}$ and $g(m) = \phi^- \mathbf{1}\{m_1 = \hat{\Theta}_{1,1}\} + \phi^+ \mathbf{1}\{m_1 = \hat{\Theta}_{1,2}\}$ with $0 \leq \phi^- < \phi^+ \leq 1$.

Definition 5: A $(2, 2)$ -mechanism is $M_1 = \{\hat{\Theta}_{1,1}, \hat{\Theta}_{1,2}\}$, $M_2 = \{\hat{\Theta}_{2,1}, \hat{\Theta}_{2,2}\}$ with $0 \leq \phi^- < \phi^+ \leq 1$, and it is either low type or high type:

$$\begin{aligned} \text{Low type: } g(m) &= \phi^- \mathbf{1}\{m = (\hat{\Theta}_{1,1}, \hat{\Theta}_{2,1})\} + \phi^+ \mathbf{1}\{\text{otherwise}\}, \\ \text{High type: } g(m) &= \phi^+ \mathbf{1}\{m = (\hat{\Theta}_{1,2}, \hat{\Theta}_{2,2})\} + \phi^- \mathbf{1}\{\text{otherwise}\}. \end{aligned}$$

A $(1, 1)$ -rule is always posterior implementable by a constant mechanism, but it reveals no information.¹⁹ The other two types of mechanisms may or may not have a posterior equilibrium. If a dictatorial mechanism for agent 1 implements a $(2, 1)$ -rule, it partially reveals agent 1's type but reveals nothing about agent 2's type. If a $(2, 2)$ -mechanism implements a $(2, 2)$ -rule, it partially reveals information to both agents. If it is a low type, the low outcome

¹⁹A constant mechanism is the only case in which, if a direct mechanism $M_1 = \Theta_1$ is used instead of $M_1 = \{\Theta_1\}$, every agent is indifferent among any messages after messages are made public. Hence, there is a posterior equilibrium in a pure strategy where type is perfectly revealed. However, once this happens, any outcome based on θ can be proposed (recall that there is no commitment). Anticipating this, a constant mechanism with full information revelation *augmented with any subsequent proposal* must have an *ex post* equilibrium. To focus on posterior implementation, we ignore this perfect information revelation in a constant mechanism.

ϕ^- needs low messages from both agents, while at least one high message results in ϕ^+ . If it is a high type, the high outcome ϕ^+ needs high messages from both agents, while at least one low message results in ϕ^- . 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 5** shows a posterior implementable low type $(3, 3)$ -rule. By **Theorem**, any posterior implementable (K, K) -rule is either a low type or a high type. The following lemma studies the existence of posterior implementable rules

Lemma 1 (existence of posterior implementable rules)

(i) *If a $(K + 1, K)$ -rule is posterior implementable for some $K \geq 2$, so is a $(2, 1)$ -rule.*

If a $(K, K + 1)$ -rule is posterior implementable for some $K \geq 2$, so is a $(1, 2)$ -rule.

If a (K, K) -rule is posterior implementable for some $K \geq 3$, so is a $(2, 2)$ -rule.

(ii) *A $(2, 1)$ -rule is posterior implementable if and only if*

$$\exists \theta_1^d \in (\underline{\theta}_1, \bar{\theta}_1) \text{ s.t. } \theta_1^d = h_1(\underline{\theta}_2, \bar{\theta}_2).$$

A $(1, 2)$ -rule is posterior implementable if and only if

$$\exists \theta_2^d \in (\underline{\theta}_2, \bar{\theta}_2) \text{ s.t. } \theta_2^d = h_2(\underline{\theta}_1, \bar{\theta}_1).$$

(iii) *A low type $(2, 2)$ -rule is posterior implementable if and only if*

$$\exists \theta^L = (\theta_1^L, \theta_2^L) \in (\underline{\theta}_1, \bar{\theta}_1) \times (\underline{\theta}_2, \bar{\theta}_2) \text{ s.t. } \theta^L = (h_1(\underline{\theta}_2, \theta_2^L), h_2(\underline{\theta}_1, \theta_1^L)).$$

A high type $(2, 2)$ -rule is posterior implementable if and only if

$$\exists \theta^H = (\theta_1^H, \theta_2^H) \in (\underline{\theta}_1, \bar{\theta}_1) \times (\underline{\theta}_2, \bar{\theta}_2) \text{ s.t. } \theta^H = (h_1(\theta_2^H, \bar{\theta}_2), h_2(\theta_1^H, \bar{\theta}_1)).$$

Lemma 1(i) shows that $(2, 1)$, $(1, 2)$, and $(2, 2)$ -rules are the key for posterior implementation. Importantly, for any information to be revealed in a posterior equilibrium, either dictatorial or $(2, 2)$ -mechanisms must have a posterior equilibrium. **Lemma 1(ii)**

and (iii) present conditions for the existence of these posterior implementable rules. Importantly, these posterior implementable rules represent endogenous information structure created by the play of a mechanism. A (2,1)-rule represents a situation where a dictator 1 can choose an outcome (ϕ^- or ϕ^+) based on his interim belief that θ_2 has a distribution $F_1(\theta_2|\theta_1, \theta_2 \in \hat{\Theta}_2)$.²⁰ Agent 2 learns about the dictator's type through his message, although she does not reveal any information as she has no choice but to accept the dictator's message.²¹ On the other hand, a (2,2)-rule gives “equal rights” to agents because each agent can choose one particular outcome independent of the other agent's message, while they need to cooperate for the other outcome.²² Also, both agents learn new information in equilibrium.²³ When a negotiation process must use information voluntarily revealed by agents, it is natural to ask when such information revelation is impossible. We turn to this issue next.

Interim agreement/disagreement. Because of **Lemma 1(i)**, ensuring the absence of dictators (i.e., no (2,1) and (1,2)-rule is posterior implementable), is necessary to prevent any information revelation. By **Lemma 1(ii)**, if neither (2,1)-rule nor (1,2)-rule is posterior implementable, there are four possible cases:

$$\begin{aligned}
A(d_1): \quad & (h_1(\underline{\theta}_2, \bar{\theta}_2), h_2(\underline{\theta}_1, \bar{\theta}_1)) = (\underline{\theta}_1, \underline{\theta}_2), \\
A(d_0): \quad & (h_1(\underline{\theta}_2, \bar{\theta}_2), h_2(\underline{\theta}_1, \bar{\theta}_1)) = (\bar{\theta}_1, \bar{\theta}_2), \\
D(1): \quad & (h_1(\underline{\theta}_2, \bar{\theta}_2), h_2(\underline{\theta}_1, \bar{\theta}_1)) = (\underline{\theta}_1, \bar{\theta}_2), \\
D(2): \quad & (h_1(\underline{\theta}_2, \bar{\theta}_2), h_2(\underline{\theta}_1, \bar{\theta}_1)) = (\bar{\theta}_1, \underline{\theta}_2).
\end{aligned} \tag{6}$$

The first case, $A(d_1)$, is where two agents agree that d_1 is better than d_0 *given their interim beliefs*: $(\theta_1, \{\theta_2 \in \Theta_2\})$ for agent 1 and $(\theta_2, \{\theta_1 \in \Theta_1\})$ for agent 2. We call it *interim agree-*

²⁰Agent 1 can choose an outcome ϕ^- by choosing $\hat{\Theta}_{1,1}$ and ϕ^+ by choosing $\hat{\Theta}_{1,2}$.

²¹All $(K+1, K)$ -rules share the property that agent 1 can choose an outcome independent of the agent 2's message.

²²For a low type (2,2)-rule, agent 1 (2) can implement ϕ^+ by choosing $\hat{\Theta}_{1,2}$ ($\hat{\Theta}_{2,2}$) regardless of the other agent's message. To implement ϕ^- , a coordinated choice of messages $(\hat{\Theta}_{1,1}, \hat{\Theta}_{2,1})$ is required.

²³All (K, K) -rules share the property that both agents can choose the same outcome independent of the other agent's message, while they need to coordinate their messages for the other outcome.

ment on d_1 . The opposite agreement pattern, $A(d_0)$, shall be called interim agreement on d_0 . The third case, $D(1)$, is where agent 1 of all types prefers d_1 to d_0 , while agent 2 of all types prefers d_0 to d_1 . The fourth case, $D(2)$, is the opposite disagreement pattern. We call $D(1)$ and $D(2)$ *interim disagreement*. Note that for all four cases, there is agreement *across types* for a given agent.

Lemma 2 (Conditions for no information revelation)

Only constant rules are posterior implementable if and only if either:

- (i) *there is interim disagreement, or*
- (ii) *there is interim agreement and there is no posterior implementable $(2, 2)$ -rule.*

Lemma 2 provides conditions which ensure no information revelation. First, interim disagreement is sufficient for no information revelation. On the other hand, when there is interim agreement, there is no $(2, 1)$ or $(1, 2)$ -rules, but information may still be revealed by implementing $(2, 2)$ -rules. In the proof of **Lemma 2**, we prove the following result.

Suppose no $(1, 2)$ and $(2, 1)$ -rule is posterior implementable.

If a low (high) type $(2, 2)$ -rule is posterior implementable,

then there is interim agreement on d_1 (on d_0).

Put differently, when coordinated messages reveal information *in the absence of dictators*, they 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 short, no information revelation is possible if and only if either (a) there is interim disagreement or (b) there is interim agreement which cannot be reversed by $(2, 2)$ -rules. **Figure 6** illustrates $A(d_1)$.

Note. In the left panel, the interim agreement on d_1 is reversed on the type set $[\underline{\theta}_1, a] \times [\underline{\theta}_2, b]$. In the right panel, there is no reversal.

Figure 6. Interim agreement on d_1 .

In **Figure 6**, a green dashed line with the end point X is a function $h_2(\underline{\theta}_1, y_1)$ for $y_1 \in \Theta_1$ and a red dashed line with the end point Y is a function $h_1(\underline{\theta}_2, y_2)$ for $y_2 \in \Theta_2$. In both panels, the flat part of h_1 implies that agent 1 with type $\underline{\theta}_1$ strictly prefers d_1 to d_0 given his interim belief. Similarly, the flat part of h_2 implies that agent 2 with type $\underline{\theta}_2$ strictly prefers d_1 to d_0 given her interim belief. Therefore, both agents strictly prefer d_1 to d_0 given their interim beliefs, i.e., interim agreement on d_1 . Given our assumptions, h_1 and h_2 are decreasing in the interior of Θ . Hence, h_1 and h_2 may or may not intersect in the interior of Θ . But when they do, the intersections of h_1 and h_2 represent posterior implementable low type (2,2)-rules by **Lemma 1(iii)**. In the left panel, the interim agreement on d_1 will be reversed to *posterior agreement* on d_0 on the posterior type set $[\underline{\theta}_1, a] \times [\underline{\theta}_2, b]$.

To see why information revelation without dictators must involve interim agreement and its posterior reversal, consider $A(d_1)$. Given his interim belief, agent 1 of all types prefers d_1 to d_0 . However, conditional on the separation of agent 2 of lower types, his preference shifts towards d_0 . Therefore, the separation of agent 2 of low type can provide incentive for agent 1 of lower types to separate, and vice versa. Notice that the separation of agent 2 is possible only if there are high types of agent 2 which will *not* switch their preference to d_0 .

after observing the separation of low type agent 1. If there were interim disagreement so that agent 2 of all types preferred d_0 to d_1 (i.e., $D(1)$ in (6)), then the separation of agent 1 of low types would only *strengthen* the interim preference of agent 2. Thus, no separation would be possible if there is interim disagreement.

Dictators v.s. coordination. Another natural question is whether a (2,2)-rule is posterior implementable when both (2,1)- and (1,2)-rules are posterior implementable. This is a situation where both agents may insist on his/her dictatorial mechanism, but a third party can suggest a (2,2)-mechanism as a compromising alternative. Such a suggestion would put both parties on more equal footing and facilitate more communication between them. 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 are posterior implementable. Let θ_1^d be the indifference type in the (2,1)-rule and θ_2^d be that in the (1,2)-rule.

(i) *If there are $k \geq 1$ posterior implementable low-type (2,2)-rules (respectively high-type), then k indifference points $\{(\theta_{1,1}^1, \theta_{2,1}^1), \dots, (\theta_{1,1}^k, \theta_{2,1}^k)\}$ lie in $(\theta_1^d, \bar{\theta}_1) \times (\theta_2^d, \bar{\theta}_2)$ (in $(\underline{\theta}_1, \theta_1^d) \times (\underline{\theta}_2, \theta_2^d)$). If they are ordered by $\theta_{1,1}^1 < \dots < \theta_{1,1}^k$, then $\theta_{2,1}^1 > \dots > \theta_{2,1}^k$.*

(ii) *Consider agent 1 with type $\bar{\theta}_1$ (respectively $\underline{\theta}_1$) believing that θ_2 is in the set $[\theta_2, \theta_2^d]$ (in $[\theta_2^d, \bar{\theta}_2]$), and agent 2 with type $\bar{\theta}_2$ ($\underline{\theta}_2$) believing that θ_1 is in the set $[\underline{\theta}_1, \theta_1^d]$ (in $[\theta_1^d, \bar{\theta}_1]$).*

(a) *If both agents strictly prefer a higher outcome or both strictly prefer a lower outcome, then k is at least one.*

(b) *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 can be zero (non-existence).*

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

Lemma 3(i) shows a trade-off between a low type (2,2)-rule and a (2,1)-rule from agent 1'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 1 can still implement ϕ^+ in the $(2, 2)$ -rule, the set of types who does so (i.e., $[\theta_1^k, \bar{\theta}_1]$ for the k th $(2, 2)$ -rule) shrinks compared with the set of such types in the $(2, 1)$ -rule ($[\theta_1^d, \bar{\theta}_1]$), because $\theta_1^d < \theta_1^k$ for all k . The same result applies to agent 2 if we compare the $(2, 2)$ -rule with a $(1, 2)$ -rule. Therefore, for each agent, the ex ante probability of choosing an outcome unilaterally is reduced in the $(2, 2)$ -rule relative to his/her dictatorial rule. In return, both agents learn new information and are given a chance to coordinate their messages.

Lemma 3(ii) provides necessary conditions for the non-existence and uniqueness of posterior implementable $(2, 2)$ -rules given that each agent can be a dictator. We illustrate cases **(a)** and **(b)** because a case **(c)** is non-generic.

Note. The left panel has three posterior implementable $(2, 2)$ -rules. The right panel has two, but the one on the left is not robust to a small change in the environment.

Figure 7. Case (a) for Lemma 3(ii).

Note. The left panel has no posterior implementable $(2, 2)$ -rule. The right panel has one but it is not robust to a small change in the environment.

Figure 8. Case (b) for Lemma 3(ii).

In **Figures 7** and **8**, a green dashed line with the boundary point X is a function $h_2(\underline{\theta}_1, y_1)$ for $y_1 \in \Theta_1$ and a red dashed line with the boundary point Y is a function $h_1(\underline{\theta}_2, y_2)$ for $y_2 \in \Theta_2$. The location of points X and Y in **Figure 7** implies that agent 1 with type $\bar{\theta}_1$ believing that $\theta_2 \in \hat{\Theta}_{2,1}$ and agent 2 with type $\bar{\theta}_2$ believing that $\theta_1 \in \hat{\Theta}_{1,1}$ both prefer d_1 to d_0 . On the other hand, the location of points X and Y in **Figure 8** implies that only the former agent prefers d_1 to d_0 . The intersections of h_1 and h_2 represent posterior implementable low type $(2, 2)$ -rules. Note that the multiple posterior implementable $(2, 2)$ -rules exist when h_1 and h_2 have multiple intersections.

At the tangent point of h_1 and h_2 (see the right panels in two figures), the intersection is not robust to a small perturbation in the environment. If we exclude this tangency case as well as case (c) with the same problem, h_1 and h_2 *must* intersect odd number of times in case (a), while they *may* intersect in case (b) even number of times. Therefore, the condition in (a) becomes necessary for $k = 1$ while the condition in (b) becomes necessary for $k = 0$. In other words, if no posterior implementable $(2, 2)$ -rule of a low-type exists, then it must be the case that the highest type agents believing that the other agent's type is below the dictator type *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.

A role of symmetry. When can we guarantee the uniqueness of a posterior implementable low type (2, 2)-rule? The slope of h_1 and h_2 represents the *expected* marginal rate of substitution of types evaluated at the points of indifference, where the expectation is taken over different type sets. When two agents are sufficiently alike ex ante, a case (a) is more likely. If, additionally, the expected marginal rate of substitution does not fluctuate much as the cutoff level changes (say h_1 and h_2 are close to *linear*), then a posterior implementable low type (2, 2)-rule is unique. On the other hand, if two agents are sufficiently different ex ante, a case like **Figure 8** is possible, and a posterior implementable low type (2, 2)-rule may not exist. The role of ex ante symmetry is formally stated below.

Definition 6: *Agents are symmetric if $\Theta_1 = \Theta_2 = [\underline{\theta}, \bar{\theta}]$, $F_1(\cdot|a) = F_2(\cdot|a) \forall a \in [\underline{\theta}, \bar{\theta}]$, and $v_1(a, b) = v_2(b, a) \forall (a, b) \in [\underline{\theta}, \bar{\theta}]^2$.*

Example 2 satisfied this symmetry condition. With symmetry, a (2, 1)-rule is posterior implementable if and only if a (1, 2)-rule is. Let the dictator's indifference type be $\theta_1^d = \theta_2^d = \theta^d \in (\underline{\theta}, \bar{\theta})$. Also, since h_1 and h_2 defined in (2) are identical with symmetry, denote both by $h : [\underline{\theta}, \bar{\theta}]^2 \rightarrow [\underline{\theta}, \bar{\theta}]^2$.

Lemma 4 (symmetric agents)

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

(i) *A posterior implementable low type (2, 2)-rule with $\theta_{1,1} = \theta_{2,1} = \theta^* \in (\theta^d, \bar{\theta})$ exists.*

Consider one agent with type $\bar{\theta}$ believing that the other agent's type is in the set $[\underline{\theta}, \theta^d]$.

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

(a) *θ^* is the only posterior implementable low type (2, 2)-rule if and only if*

$$\forall y < \theta^*, h(\underline{\theta}, h(\underline{\theta}, y)) < y.$$

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

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

(a) θ^* is the only posterior implementable low type $(2, 2)$ -rule if and only if

$$\forall y > \theta^*, h(\underline{\theta}, h(\underline{\theta}, y)) < y.$$

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

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

(a) θ^* is the only posterior implementable low type $(2, 2)$ -rule if and only if

$$\forall y \in (\theta^*, \bar{\theta}) \ h(\underline{\theta}, h(\underline{\theta}, y)) \neq y.$$

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

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

(v) A posterior implementable $(2, 2)$ -rule exists if and only if $\{\theta \in \Theta | v_1(\theta) = 0\} \cap (\underline{\theta}, \bar{\theta})^2$ is not empty.

An analogous result for a high type $(2, 2)$ -rule holds but is not presented. From the converse of **Lemma 4(i)**, given symmetric agents, if there is no symmetric posterior implementable $(2, 2)$ -rule, then there is no posterior implementable $(1, 2)$ - or $(2, 1)$ -rule. By **Lemma 1**, this implies that only constant rules are posterior implementable. Hence, with symmetry, information revelation occurs if and only if $(2, 2)$ -rules are posterior implementable. **Lemma 4(v)** provides the condition for the absence of posterior implementable $(2, 2)$ -rules. With symmetry, information revelation is impossible if and only if $v_1(\theta)$ (hence $v_2(\theta)$) does not go through the interior of the type set Θ . This is the case where agents prefer the same decision in all states.²⁴ Except this trivial situation, posterior implementable $(2, 2)$ -rules always exist if agents are symmetric.

5 Posterior Renegotiation-proof Decision Rules

Because a posterior implementable (K, L) -rule ϕ can realize only one of the $K \times L$ posterior type sets in $\Theta_\phi = \left\{ \hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l} \right\}_{k=1, \dots, K, l=1, \dots, L}$, for each $\hat{\Theta}_{k,l} = \hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l} \in \Theta_\phi$, potential renegotiation is represented by posterior implementable rules defined on each $\hat{\Theta}_{k,l}$. For each

²⁴With the possible indifference for the lowest (highest) type observing the lowest (highest) type.

$\hat{\Theta}_{k,l} \in \Theta_\phi$, the relevant status quo outcome is $\phi(\hat{\Theta}_{k,l})$. Importantly, Assumptions 1, 2, 3 continue to hold in each posterior type set $\hat{\Theta}_{k,l} \in \Theta_\phi$.²⁵ Therefore, all the properties of posterior implementable rules defined in the original type set Θ apply to the posterior implementable rules defined in each posterior type set $\hat{\Theta}_{k,l}$.

Definition 7: Let ϕ be a posterior implementable (K, L) -rule with a partition Θ_ϕ . Say ϕ is NOT posterior renegotiation-proof if $\exists \hat{\Theta}_{k,l} \in \Theta_\phi$ s.t. there is a rule $\hat{\phi}$ defined on $\hat{\Theta}_{k,l}$ with one of the following two properties:

- (a) $\hat{\phi}$ is posterior implementable and not constant, or
- (b) $\hat{\phi}$ is constant and both agents of all types in $\hat{\Theta}_{k,l}$ weakly prefer $\hat{\phi}$ to $\phi(\hat{\Theta}_{k,l})$ and at least one agent of some types in $\hat{\Theta}_{k,l}$ strictly prefers $\hat{\phi}$ to $\phi(\hat{\Theta}_{k,l})$.

Posterior renegotiation-proofness is defined by the absence of both “posterior information revelation” and “unanimous improvement *without* information revelation”. In particular, (a) in the definition requires that, in any posterior type set $\hat{\Theta}_{k,l}$, there is no posterior equilibrium in which information revelation occurs.²⁶ 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 for all types, 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) eliminates this possibility.

In the above definition, the second rule $\hat{\phi}$ must be posterior implementable, but does not have to pass the exactly same test applied to the first rule ϕ . We do not impose such

²⁵Assumption 1 and 2(a) hold in any $\hat{\Theta}_{k,l}$. Because Assumptions 2(b) and 3 were made for *any* subintervals $\hat{\Theta}_j$ of Θ_j , they also hold in any $\hat{\Theta}_{k,l}$.

²⁶This is similar to strong renegotiation-proofness in Maskin and Tirole (1992).

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 solution concept seems reasonable in the following sense. Suppose that a firm finds a profitable opportunity (i.e., a mechanism that will be accepted by agents relative to the status quo) but is aware that with some probability (i.e., for some types of agents) a new competitor might steal its business later. We argue 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 potential 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 no low type $(2, 2)$ -rule is posterior implementable,*
 - (c) *interim agreement on d_0 , $\phi^0 = 0$, and no high type $(2, 2)$ -rule is posterior implementable.*
- (iii) *A posterior implementable $(2, 1)$ -rule with $\theta_1^d \in (\underline{\theta}_1, \bar{\theta}_1)$ and $\phi^- < \phi^+$ is PRP if and only if one of the following conditions (a)-(c) holds:*
 - (a) $h_2(\underline{\theta}_1, \theta_1^d) = \underline{\theta}_2$ and $\phi^+ = 1$,
 - (b) $h_2(\theta_1^d, \bar{\theta}_1) = \bar{\theta}_2$ and $\phi^- = 0$,
 - (c) $h_2(\underline{\theta}_1, \theta_1^d) = \bar{\theta}_2$, $h_2(\theta_1^d, \bar{\theta}_1) = \underline{\theta}_2$, $\phi^- = 0$, $\phi^+ = 1$,
no high type $(2, 2)$ -rule is posterior implementable in the set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$,
no low type $(2, 2)$ -rule is posterior implementable in the set $[\theta_1^d, \bar{\theta}_1] \times \Theta_2$.
- (iv) *A posterior implementable low type $(2, 2)$ -rule with $\phi^- < \phi^+$ is PRP*

if and only if $\phi^- = 0$, $\phi^+ = 1$ and all of the following conditions are satisfied:

- (a) no $(1, 2)$ -rule and low type $(2, 2)$ -rule is posterior implementable on $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$,
 - (b) no $(2, 1)$ -rule and low type $(2, 2)$ -rule is posterior implementable on $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,2}$,
 - (c) no $(2, 2)$ -rule of a high type is posterior implementable on $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$.
- (v) (a) A posterior implementable $(3, 2)$ -rule with $\phi^- < \phi^+$ is PRP if and only if $\phi^- = 0$, $\phi^+ = 1$, and no low type $(2, 2)$ -rule is posterior implementable on $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,2}$ and no high type $(2, 2)$ -rule is posterior implementable on $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$.
- (b) A $(3, 2)$ -rule is PRP only if $\{\theta \in \Theta | v_2(\theta) = 0\} \subseteq \hat{\Theta}_{1,2} \times \Theta_2$.

Proposition characterizes PRP rules.²⁷ The first part (i) reveals a key feature of PRP rules that they cannot have more than five messages. Considering that there is a continuum of types, and also that there can be posterior implementable rules with infinite number of partitions²⁸, the result shows that information revelation is significantly reduced without commitment to a mechanism. This is because whenever two agents have revealed that they are “middle” types, it is possible to make both agents reveal more information.²⁹

Note. $h_1(b, y_2)$ and $h_2(a, y_1)$ are defined in the posterior type set $[a, a2] \times [b, b2]$. The

²⁷(iii) and (v) have symmetric counterparts for $(1, 2)$ - and $(2, 3)$ -rules and (iv) has a counterpart for high type, but they are omitted.

²⁸See Green and Laffont (1987) page 84 for a discussion of an accumulation point.

²⁹In fact, it is always possible to make both agents better off when they reveal that they are middle types. See the discussion at the end of Example 1.

intersection of h_1 and h_2 (a blue dot) indicates a posterior implementable low type $(2, 2)$ -rule defined in the type set $[a, a2] \times [b, b2]$.

Figure 9. Posterior implementable $(3, 3)$ -rules invite renegotiation.

Figure 9 shows a posterior implementable low type $(3, 3)$ -rule. Recall that the step function consists of vertical segments (associated with indifference types of agent 1) and horizontal segments (associated with indifference types of agent 2). Therefore, $v_1(\theta) = 0$ must go through the interior of vertical segments (see red dots), while $v_2(\theta) = 0$ must go through the interior of horizontal segments (see green dots). This implies that $v_1(\theta) = 0$ must also go through the interior of any horizontal segments between two vertical segments. Similarly, $v_2(\theta) = 0$ must also go through the interior of any vertical segments between two horizontal segments. Therefore, for any posterior implementable (K, L) -rules with $K + L \geq 5$, both $v_1(\theta) = 0$ and $v_2(\theta) = 0$ must go through any segments of the step function *except two segments at the both ends* of the step function. If $K + L \geq 6$, there is at least one posterior type set in the interior of Θ whose boundaries are two segments of the step function (see the set $[a, a2] \times [b, b2]$ in **Figure 9**). By construction, $h_1(b, y_2)$ and $h_2(a, y_1)$ defined on this set must intersect at least once in the interior.

Proposition(ii)-(v) identifies eight different PRP rules (ignoring omitted counterparts for (iii)-(v)). To understand the qualitative nature of different PRP rules, recall the interim agreement/disagreement defined in (6). We can similarly define *posterior agreement/disagreement* for each posterior type set created by posterior implementable rules. For example, there is posterior agreement on the posterior type set $\hat{\Theta}_{k,l} \in \Theta_\phi$ if two agents of all types agree that one decision is better than the other decision *given their posterior beliefs*: $(\theta_1, \{\theta_2 \in \hat{\Theta}_{2,l}\})$ for agent 1 and $(\theta_2, \{\theta_1 \in \hat{\Theta}_{1,k}\})$ for agent 2. If ϕ is a PRP rule, by **Lemma 2** applied to each posterior type set, there must be either posterior agreement or posterior disagreement in each $\hat{\Theta}_{k,l} \in \Theta_\phi$. This leads to the following classification of **Proposition(ii)-(v)** based on posterior agreement/disagreement.

For (ii)(a) and (iii)(a)(b), there is posterior disagreement in some posterior type sets.

For (ii)(b)(c), (iii)(c), (iv), (v), there is posterior agreement in each posterior type set.

PRP rules with posterior disagreement ((ii)(a) and (iii)(a)(b)) have a property that an intermediate outcome strictly between zero and one (i.e., random mixture of d_0 and d_1) can be implemented.³⁰ In any posterior type set $\hat{\Theta}_{k,l}$ with posterior disagreement, by **Lemma 2(i)**, constant rules are the only posterior implementable rules. Moreover, any change to $\phi^0(\hat{\Theta}_{k,l})$ would be vetoed by one agent of *all types*, hence *without revealing any information*. These PRP rules are shown in **Figure 10**. The remaining PRP rules ((ii)(a) and (iii)(a)(b)) have a property that posterior agreement occurs in each posterior type set. Therefore, these rules must have $\phi^- = 0$ and $\phi^+ = 1$ to avoid improvement by constant rules. Additionally, these cases require that no (2, 2)-rule is posterior implementable. These PRP rules are shown in **Figures 11** and **12**.

Note. The upper panels show Proposition(ii)(a). The lower left panel shows (iii)(a) and the lower right panel shows (iii)(b). For illustration, $\underline{\theta}_1 = \underline{\theta}_2 = 0$ and $\bar{\theta}_1 = \bar{\theta}_2 = 1$ are assumed.

Figure 10. PRP rules with posterior disagreement.

³⁰In (ii)(a) any $\phi^0 \in [0, 1]$ is possible, in (iii)(a) any $\phi^- < 1$ is possible, and in (iii)(b) any $\phi^+ > 0$ is possible.

We explain only the lower-left panel of **Figure 10**, which corresponds to **(iii)(a)** in **Proposition**. There is posterior disagreement on the posterior type set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$, where agent 2 of all types prefers d_1 to d_0 and agent 1 of all types prefers d_0 to d_1 . By **Lemma 2(i)** applied in the set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$, only constant rules can challenge the status quo. However, due to posterior disagreement, any change to ϕ^- would be vetoed by one agent *without revealing any information*.³¹ Note that $h_2(\underline{\theta}_1, \theta_1^d) = \underline{\theta}_2$ implies that the solution to $h_2(\underline{\theta}_1, y_1) = \underline{\theta}_2$ (point X in the figure) is no larger than θ_1^d . This implies two things. First, there is posterior disagreement (only agent 2 prefers d_1 , i.e., $D(2)$) on the posterior type set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$. Second, there is posterior agreement on d_1 in the posterior type set $[\theta_1^d, \bar{\theta}_1] \times \Theta_2$, and $h_1(\underline{\theta}_2, y_2)$ and $h_2(\underline{\theta}_1, y_1)$ do not intersect in this set. Thus, this $(2, 1)$ -rule is PRP.

Note. The left two panels show (ii)(b) and (ii)(c) in Proposition. The right panel shows (iii)(c) with $\theta_1^d = a$. For illustration, $\underline{\theta}_1 = \underline{\theta}_2 = 0$ and $\bar{\theta}_1 = \bar{\theta}_2 = 1$ are assumed.

Figure 11.

Figure 11 illustrates PRP rules in **Proposition(ii)(b)(c)** and **(iii)(c)**. In each panel, the location of X and Y indicates that no $(1, 2)$ and $(2, 1)$ -rule is posterior implementable in each posterior type set. Additionally, arrows indicate that no $(2, 2)$ -rule is posterior implementable in each posterior type set.

³¹In **(iii)(b)**, the same explanation applies to posterior disagreement in the posterior type set $[\theta_1^d, \bar{\theta}_1] \times \Theta_2$. In **(ii)(a)**, there is interim disagreement and the same explanation applies in the type set Θ .

Note. The left panel is a PRP (2, 2)-rule of low type. The right panel is a PRP (3, 2)-rule.
 h_1 and h_2 are defined in each posterior type set.

Figure 12. A posterior renegotiation-proof (2, 2)- and (3, 2)-rules.

Figure 12 illustrates **Proposition(iv) and (v)**. The left panel shows a PRP (2, 2)-rule of low type. There are four posterior type sets A, B, C, D created by this rule. The rule assigns the low outcome ϕ^- in the set B and the high outcome ϕ^+ in the other sets A, C, D . Because $\theta_{2,1}$ is an indifference type of agent 2 believing that agent 1's type is in the set $[\underline{\theta}_1, \theta_{1,1}]$, $v_2(\theta) = 0$ must go through the interior of the boundary line between A and B (see a green dot). Similarly, $v_1(\theta) = 0$ must go through the interior of the boundary line between B and D (see a red dot).

In the set C , there is ex post agreement on d_1 . Therefore, if $\phi^+ < 1$, improvement with any higher outcome is possible in C . Thus, ϕ^+ must be one. In the set A , under complete information, agents would disagree in the area left to $v_2(\theta) = 0$ (only agent 1 prefers d_1) and they agree on d_1 in the other area. However, the information revealed in the first mechanism creates posterior agreement on d_1 in A , because agent 2 of all types $\theta_2 \in [\theta_{2,1}, \bar{\theta}_2]$ weakly prefers d_1 believing that agent 1's type is in the set $[\underline{\theta}_1, \theta_{1,1}]$. Moreover, no more information revelation is possible because no (1, 2), (2, 1), and (2, 2)-rule is posterior implementable in A .³² Because $\phi^+ = 1$, the set A passes the test of PRP.

³²First, no posterior implementable (1, 2)-rule exists above $\theta_{2,1}$. Also, figure shows that $h_1(\theta_{2,1}, y_2) = \underline{\theta}_1 \forall y \in [\theta_{2,1}, \bar{\theta}_2]$. This implies that no (2, 1) and (2, 2)-rule is posterior implementable in A .

In the set D , under complete information, two agents agree on d_0 in the area below $v_2(\theta) = 0$, they agree on d_1 in the area above $v_1(\theta) = 0$, and they disagree in the area between two curves. However, there is posterior agreement on d_1 conditional on the information revealed by the first mechanism, and no more information revelation is possible, as indicated by the location of point X and the arrow. With $\phi^+ = 1$, the area D also passes the test of PRP. Finally, in the set B , under complete information both agreement and disagreement are possible. However, there is posterior agreement on d_0 conditional on the information revealed by the first mechanism, and no more information revelation is possible, as indicated by the arrow. With $\phi^- = 0$, the area B also passes the test of PRP. All in all, the posterior implementable $(2, 2)$ -rule in **Figure 12** is a PRP rule.

The mechanism that implements this rule is simple (two messages for each agent) and may appear to leave much room for potential renegotiation. For example, there are states in B and D where both agents would oppose to the mechanism's proposal, *if they can reveal these states*. However, there is posterior 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 voluntarily revealed. We leave a similar analysis of $(3, 2)$ -rules for interested readers.

Symmetry revisited. 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 for $(2, 2)$ -rules.

Corollary (PRP with symmetric agents)

*Suppose that agents are symmetric in the sense of **Definition 6**.*

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

- (a) $v_1(\theta) \geq 0 \ \forall \theta \in [\underline{\theta}, \bar{\theta}]^2$ and $\phi^0 = 1$, or
- (b) $v_1(\theta) \leq 0 \ \forall \theta \in [\underline{\theta}, \bar{\theta}]^2$ and $\phi^0 = 0$.

- (ii) A posterior implementable $(2, 1)$ -rule with $\theta_1^d \in (\underline{\theta}, \bar{\theta})$ and $\phi^- < \phi^+$ is PRP if and only if **Proposition (iii)(c)** holds.
- (iii) A posterior implementable $(2, 2)$ -rule with indifferent types $\theta_{1,1}$ and $\theta_{2,1}$ is PRP only if $\theta_{1,1} \neq \theta_{2,1}$.

Compared with **Proposition**, cases with posterior disagreement do not arise with symmetry. Put differently, the model with symmetric agents cannot describe a situation where agents disagree on the final outcome. **Corollary(iii)** allows us to construct an example where no PRP rule exists. By **Lemma 4(i)(v)**, with symmetry, a $(2, 2)$ -rule exists except a trivial case of no conflict of interests. However, if the $(2, 2)$ -rule is symmetric in the sense of $\theta_{1,1} = \theta_{2,1}$ (which always exists with symmetry when any $(2, 2)$ -rule is posterior implementable), then two of the posterior type sets $[\underline{\theta}, \theta_{1,1}] \times [\underline{\theta}, \theta_{2,1}]$ and $[\theta_{1,1}, \bar{\theta}] \times [\theta_{2,1}, \bar{\theta}]$ create symmetric environments again. Because v_1 and v_2 intersect in at least one of these sets, by **Lemma 4(v)**, there is a posterior implementable $(2, 2)$ -rule. Because a symmetric posterior implementable $(2, 2)$ -rule always has a posterior type set with renegotiation, if such a rule is a unique non-constant posterior implementable rule, no PRP rule exists.³³ Even ex ante symmetric agents must *look asymmetric* if a PRP rule is implemented.

6 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 requiring renegotiation-proofness puts significant restrictions on the amount of information revelation.

³³For example, consider a modified version of Example 1 where $c_1 = c_2 = c$ is sufficiently small. The only posterior implementable rule which is not constant is a low type $(2, 2)$ -rule, but it is not PRP. In sum, if (i) agents are symmetric, (ii) no $(2, 1)$ -rule is posterior implementable, and (iii) a unique posterior implementable $(2, 2)$ -rule has $\theta_{1,1} = \theta_{2,1}$, then no PRP rule exists.

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 messages 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.³⁴ 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.

7 Proofs

Proof of Lemma 1.

(i) First, consider a $(K + 1, K)$ -rule. If it is posterior implementable, by Theorem, there exists $(\theta_{1,1}, \theta_{1,K})$ such that $\underline{\theta}_1 < \theta_{1,1} = h_1(\theta_{2,K-1}, \bar{\theta}_2)$ and $\theta_{1,K} = h_1(\underline{\theta}_2, \theta_{2,1}) < \bar{\theta}_1$. Because h_1 is strictly decreasing in both arguments, $\theta_{1,1} = h_1(\theta_{2,K-1}, \bar{\theta}_2) < h_1(\underline{\theta}_2, \bar{\theta}_2) < h_1(\underline{\theta}_2, \theta_{2,1}) = \theta_{1,K}$. Because $\theta_1^d \equiv h_1(\underline{\theta}_2, \bar{\theta}_2)$ is an indifferent type of 1 believing that agent 2's type is in Θ_2 , a $(2, 1)$ -mechanism with $\hat{\Theta}_{1,1} = [\underline{\theta}_1, \theta_1^d]$ and $\hat{\Theta}_{1,2} = [\theta_1^d, \bar{\theta}_1]$ has a posterior equilibrium which implements a $(2, 1)$ -rule. The proof is symmetric for a $(K, K + 1)$ -rule.

Next, consider a (K, K) -rule of a low type with $K \geq 3$ (as in **Figure 5**). The proof for a high type is similar and omitted. If it is posterior implementable, by Theorem, there exists $(\theta_{1,1}, \theta_{1,K-1})$ such that $\underline{\theta}_1 < \theta_{1,1} = h_1(\theta_{2,K-2}, \theta_{2,K-1}) < \theta_{1,K-1} = h_1(\underline{\theta}_2, \theta_{2,1}) < \bar{\theta}_1$. Similarly, there exists $(\theta_{2,1}, \theta_{2,K-1})$ such that $\underline{\theta}_2 < \theta_{2,1} = h_2(\theta_{1,K-2}, \theta_{1,K-1}) < \theta_{2,K-1} = h_2(\underline{\theta}_1, \theta_{1,1}) < \bar{\theta}_2$. Let $X_1 \equiv \bigcup_{k=2}^{K-1} \hat{\Theta}_{1,k}$ and $X_2 \equiv \bigcup_{k=2}^{K-1} \hat{\Theta}_{2,k}$. Because h_1 and h_2 are strictly decreasing in both arguments, $\min X_1 = \theta_{1,1} = h_1(\theta_{2,K-2}, \theta_{2,K-1}) < h_1(\underline{\theta}_2, \theta_{2,K-1}) < h_1(\underline{\theta}_2, \theta_{2,1}) = \theta_{1,K-1} = \max X_1$ and $\min X_2 = \theta_{2,1} = h_2(\theta_{1,K-2}, \theta_{1,K-1}) < h_2(\underline{\theta}_1, \theta_{1,K-1}) < h_2(\underline{\theta}_1, \theta_{1,1}) = \theta_{2,K-1} = \max X_2$. Therefore, $(x_1^*, x_2^*) \equiv (h_1(\underline{\theta}_2, \theta_{2,K-1}), h_2(\underline{\theta}_1, \theta_{1,K-1}))$ lies in the interior of $X_1 \times X_2$. Define $h_1 : X_2 \rightarrow X_1$ by $h_1(\underline{\theta}_2, x_2) \forall x_2 \in X_2$. Similarly define $h_2 : X_1 \rightarrow X_2$ by $h_2(\underline{\theta}_1, x_1)$

³⁴Watson (1999) is a related work in a dynamic environment.

$\forall x_1 \in X_1$. Note that $h_1(\cdot)$ and $h_2(\cdot)$ are continuous, strictly decreasing, $h_1(\theta_{2,1}) = \theta_{1,K-1}$, $h_2(\theta_{1,1}) = \theta_{2,K-1}$, $h_1(\theta_{2,K-1}) = x_1^*$, and $h_2(\theta_{1,K-1}) = x_2^*$. Hence, a mapping $t : X_1 \times X_2 \rightarrow X_1 \times X_2$ defined by $t(x_1, x_2) = (h_1(x_2), h_2(x_1))$ has at least one fixed point in the set $(x_1^*, \theta_{1,K-1}) \times (x_2^*, \theta_{2,K-1})$. This fixed point constitutes a low type (2, 2)-rule.

(ii) We prove for a (2, 1)-rule. The proof for a (1, 2)-rule is similar and omitted.

(if) By Assumptions 1 through 3, $\theta_1 \leq \theta_1^d \Leftrightarrow V_1(\theta_1; [\underline{\theta}_2, \bar{\theta}_2]) \leq 0$. Therefore, agent 1 with $\theta_1 < \theta_1^d$ prefers a higher outcome while agent 1 with $\theta_1 > \theta_1^d$ prefers a lower outcome. Therefore, a dictatorial mechanism where a message $\hat{\Theta}_{1,1} = [\underline{\theta}_1, \theta_1^d]$ achieves ϕ^- and a message $\hat{\Theta}_{1,2} = [\theta_1^d, \bar{\theta}_1]$ achieves $\phi^+ > \phi^-$ has a posterior equilibrium which results in a (2, 1)-rule.

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

$$\begin{aligned} \forall \theta_1 \in \Theta_1, V_1(\theta_1; [\underline{\theta}_2, \bar{\theta}_2]) &\geq 0 \Leftrightarrow h_1(\underline{\theta}_2, \bar{\theta}_2) = \underline{\theta}_1 \text{ or} \\ \forall \theta_1 \in \Theta_1, V_1(\theta_1; [\underline{\theta}_2, \bar{\theta}_2]) &\leq 0 \Leftrightarrow h_1(\underline{\theta}_2, \bar{\theta}_2) = \bar{\theta}_1. \end{aligned}$$

For the former case, agent 1 always prefers a higher outcome and $\hat{\Theta}_{1,1}$ would not be chosen. For the latter case, agent 1 always prefers a lower outcome and $\hat{\Theta}_{1,2}$ would not be chosen. Therefore, a (2, 1)-rule is not posterior implementable.

(iii) We prove for a low type (2, 2)-rule. The proof for a high type is similar and omitted.

(if) Consider a mechanism $M = \{L, H\}^2$, $g(L, L) = \phi^-$, $g(L, H) = g(H, L) = g(H, H) = \phi^+$. We verify the optimality of the following strategies: agent 1 chooses L if his type is below θ_1^L and chooses H if his type is above θ_1^L , agent 2 chooses L if his type is below θ_2^L and chooses H if his type is above θ_2^L . Given agent 2's strategy, if agent 1 with type θ_1 observes agent 2's message L , his expected payoff is $V_1(\theta_1; [\underline{\theta}_2, \theta_2^L])$. By the definition of $\theta_1^L = (h_1(\underline{\theta}_2, \theta_2^L), h_2(\theta_1, \theta_1^L))$, $\theta_1 \leq \theta_1^L \Leftrightarrow V_1(\theta_1; [\underline{\theta}_2, \theta_2^L]) \leq 0$. Hence, agent 1 with $\theta_1 < \theta_1^L$ prefers a lower outcome and choosing L is optimal, while agent 1 with $\theta_1 > \theta_1^L$ prefers a higher outcome and choosing H is optimal. If agent 1 observes agent 2's message H , the outcome ϕ^+ is independent of his message. Therefore, agent 1's strategy is optimal. The optimality of agent 2's strategy can be similarly verified.

(only if) If the condition is not satisfied, $h_1(\underline{\theta}_2, y_2)$ and $h_2(\underline{\theta}_1, y_1)$ do not share a point in the interior of Θ . Hence, we need to consider the six cases:

- (a) $h_1(\underline{\theta}_2, y_2) = \underline{\theta}_1$ for all $y_2 \in (\underline{\theta}_2, \bar{\theta}_2)$,
- (b) $h_1(\underline{\theta}_2, y_2) = \bar{\theta}_1$ for all $y_2 \in (\underline{\theta}_2, \bar{\theta}_2)$,
- (c) $\forall y_2 \in (\underline{\theta}_2, \bar{\theta}_2)$ s.t. $h_1(\underline{\theta}_2, y_2) \in (\underline{\theta}_1, \bar{\theta}_1)$, $V_2(y_2; [\underline{\theta}_1, h_1(\underline{\theta}_2, y_2)]) \neq 0$.

Cases (d), (e), (f) are obtained by switching the role of 1 and 2 in (a)-(c). First note that to implement a low type (2, 2)-rule, agent 1 must prefer a low outcome if his type is below some threshold, and a high outcome if his type is above it, given the belief that agent 2's type is in some set $[\underline{\theta}_2, y_2]$ with $y_2 \in (\underline{\theta}_2, \bar{\theta}_2)$. For (a) and (b), agent 1 always prefers one outcome to the other (ϕ^+ for (a) and ϕ^- for (b)) given the belief that agent 2's type is in the set $[\underline{\theta}_2, y_2]$ for any $y_2 \in (\underline{\theta}_2, \bar{\theta}_2)$. Hence, it is impossible to implement a low type (2, 2)-rule. For (c), suppose that a low type (2, 2)-rule is posterior implementable. Then there must be $y_2 \in (\underline{\theta}_2, \bar{\theta}_2)$ such that $h_1(\underline{\theta}_2, y_2) \in (\underline{\theta}_1, \bar{\theta}_1)$ and, given the belief that agent 1's type is below

$h_1(\underline{\theta}_2, y_2)$, agent 2 with type above y_2 prefers a high outcome while agent 2 with type below y_2 prefers a low outcome. This requires $V_2(\theta_2; [\underline{\theta}_1, h_1(\underline{\theta}_2, y_2)]) \leq 0 \Leftrightarrow \theta_2 \leq y_2$. By continuity of V_2 , this implies $V_2(y_2; [\underline{\theta}_1, h_1(\underline{\theta}_2, y_2)]) = 0$ and contradicts (c). The remaining cases are symmetric. ■(L1)

Proof of Lemma 2.

We prove the following result which directly proves Lemma 2.

Only constant rules are posterior implementable if and only if one of the following holds:

- (i) *there is interim disagreement,*
- (ii) *there is interim agreement on d_0 and there is no posterior implementable high type (2, 2)-rule,*
- (iii) *there is interim agreement on d_1 and there is no posterior implementable low type (2, 2)-rule.*

First, we show that no (2, 1) and (1, 2)-rule is posterior implementable if and only if one of $A(d_0)$, $A(d_1)$, $D(1)$, $D(2)$ is true. By Lemma 1(ii), no (2, 1)-rule is posterior implementable if and only if $h_1(\underline{\theta}_2, \bar{\theta}_2) = \underline{\theta}_1$ or $h_1(\underline{\theta}_2, \bar{\theta}_2) = \bar{\theta}_1$. Similarly, no (1, 2)-rule is posterior implementable if and only if $h_2(\underline{\theta}_1, \bar{\theta}_1) = \underline{\theta}_2$ or $h_2(\underline{\theta}_1, \bar{\theta}_1) = \bar{\theta}_2$. Hence, there are four possible combinations. First, $h_1(\underline{\theta}_2, \bar{\theta}_2) = \underline{\theta}_1$ and $h_2(\underline{\theta}_1, \bar{\theta}_1) = \underline{\theta}_2$ imply that both agents prefer d_1 to d_0 given their interim beliefs ($A(d_1)$), while $h_1(\underline{\theta}_2, \bar{\theta}_2) = \bar{\theta}_1$ and $h_2(\underline{\theta}_1, \bar{\theta}_1) = \bar{\theta}_2$ imply that both agents prefer d_0 to d_1 given their interim beliefs ($A(d_0)$). Next, $h_1(\underline{\theta}_2, \bar{\theta}_2) = \underline{\theta}_1$ and $h_2(\underline{\theta}_1, \bar{\theta}_1) = \bar{\theta}_2$ imply that agent 1 prefers d_1 to d_0 , while agent 2 prefers d_0 to d_1 given their interim beliefs ($D(1)$). Finally, $h_1(\underline{\theta}_2, \bar{\theta}_2) = \bar{\theta}_1$ and $h_2(\underline{\theta}_1, \bar{\theta}_1) = \underline{\theta}_2$ imply that agent 1 prefers d_0 to d_1 , while agent 2 prefers d_1 to d_0 given their interim beliefs ($D(2)$).

(if)

(i) By Lemma 1(i), it suffices to show that (2, 2)-rules are not posterior implementable if interim disagreement exists. Consider $D(1)$. Because $h_1(x_2, \bar{\theta}_2) \leq h_1(\underline{\theta}_2, \bar{\theta}_2) = \underline{\theta}_1 \forall x_2 \in \Theta_2$, $h_1(x_2, \bar{\theta}_2)$ and $h_2(x_1, \bar{\theta}_1)$ cannot share a point in the interior of Θ . Hence, a high type (2, 2)-rule is not posterior implementable. Similarly, a low type (2, 2)-rule is not posterior implementable because $h_2(\underline{\theta}_1, x_1) \geq h_2(\underline{\theta}_1, \bar{\theta}_1) = \bar{\theta}_2 \forall x_1 \in \Theta_1$. The case $D(2)$ is similar.

(ii) Because interim agreement implies that no (2, 1) and (1, 2)-rule is posterior implementable, it suffices to show that interim agreement on d_0 implies that there is no posterior implementable low type (2, 2)-rule. We prove this by showing the converse. Suppose that a low type (2, 2)-rule is posterior implementable and there is $A(d_0)$. This implies $\bar{\theta}_1 = h_1(\underline{\theta}_2, \bar{\theta}_2)$. Let $\theta_{1,1}$ and $\theta_{2,1}$ be indifferent types for a low type (2, 2)-rule. Then $\theta_{1,1} = h_1(\underline{\theta}_2, \theta_{2,1}) < \bar{\theta}_1 = h_1(\underline{\theta}_2, \bar{\theta}_2)$. This contradicts that h_1 is decreasing in the second argument. Hence, if a low type (2, 2)-rule is posterior implementable in the presence of interim agreement, it must be $A(d_1)$.

(iii) The proof is symmetric to (ii) and omitted.

(Only if)

If only constant rules are posterior implementable, then no (2, 1) and (1, 2)-rule is posterior implementable and one of $A(d_0)$, $A(d_1)$, $D(1)$, $D(2)$ must be true. We already showed that (2, 2)-rules are not posterior implementable if interim disagreement exists. Because we also showed that $A(d_1)$ is compatible only with a low type (2, 2)-rule and that $A(d_0)$

is compatible only with a high type (2,2)-rule. Therefore, (i)-(iii) exhaust all the possible cases. ■(L2)

Proof of Lemma 3.

We prove for a low type (2,2)-rule. The proof for a high type is similar and omitted.

(i) Define $h_1 : \Theta_2 \rightarrow \Theta_1$ by $h_1(\underline{\theta}_2, y_2)$ and $h_2 : \Theta_1 \rightarrow \Theta_2$ by $h_2(\underline{\theta}_1, y_1)$. Then $h_1(\underline{\theta}_2, \bar{\theta}_2) = \theta_1^d$ and h_1 is strictly decreasing. Similarly, $h_2(\underline{\theta}_1, \bar{\theta}_1) = \theta_2^d$ and h_2 is strictly decreasing. By **Lemma 1**, a low type (2,2)-rules exist if and only if h_1 and h_2 have intersections. Because both h_1 and h_2 are strictly decreasing, all the intersections must lie in the set $(\theta_1^d, \bar{\theta}_1) \times (\theta_2^d, \bar{\theta}_2)$ and ordered as stated.

(ii) (a) The stated condition implies either $h_1(\underline{\theta}_2, \theta_2^d) < \bar{\theta}_1$ and $h_2(\underline{\theta}_1, \theta_1^d) < \bar{\theta}_2$ (both agents in question prefer a higher outcome. This is **Figure 7**) or $\exists \theta_1 > \theta_1^d$ s.t. $h_2(\underline{\theta}_1, \theta_1) = \bar{\theta}_2$ and $\exists \theta_2 > \theta_2^d$ s.t. $h_1(\underline{\theta}_2, \theta_2) = \bar{\theta}_1$ (both prefer a lower outcome). Either way, because h_1 and h_2 are strictly decreasing, they must cross each other at least once in $[\theta_1^d, \bar{\theta}_1] \times [\theta_2^d, \bar{\theta}_2]$.

(b) The stated condition implies either $h_1(\underline{\theta}_2, \theta_2^d) \leq \bar{\theta}_1$ and $\exists \theta_1 \geq \theta_1^d$ s.t. $h_2(\underline{\theta}_1, \theta_1) = \bar{\theta}_2$ and at least one inequality is strict (only agent 1 in question prefers a higher outcome. This is **Figure 8**), or $h_2(\underline{\theta}_1, \theta_1^d) \leq \bar{\theta}_2$ and $\exists \theta_2 \geq \theta_2^d$ s.t. $h_1(\underline{\theta}_2, \theta_2) = \bar{\theta}_1$ and at least one inequality is strict (only agent 2 prefers a higher outcome). Either way, there are three possibility: in the interior of $[\theta_1^d, \bar{\theta}_1] \times [\theta_2^d, \bar{\theta}_2]$, 1) h_1 and h_2 do not share a point, 2) they cross each other for even number of times, 3) they have tangent points. Note that $k = 1$ is possible only if there is a unique tangent point and no crossing. Such a case is not robust to a small change in the environment.

(c) The stated condition implies that h_1 and h_2 intersect at two points $(\theta_1^d, \bar{\theta}_2)$ and $(\bar{\theta}_1, \theta_2^d)$, but these two points do not constitute posterior implementable (2,2)-rules. Two strictly decreasing curves sharing the same end points can intersect in the middle any number of times. ■(L3)

Proof of Lemma 4.

(i) When agents are symmetric, case (ii)(b) in **Lemma 3** cannot happen because h_1 and h_2 are located symmetrically with respect to a straight line connecting two points $(\underline{\theta}, \underline{\theta})$ and $(\bar{\theta}, \bar{\theta})$ on $[\underline{\theta}, \bar{\theta}]^2$. Call this line the 45 degree line. Because h_1 and h_2 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) $h(\underline{\theta}, \theta_2)$ starts from the interior of a segment $\{\bar{\theta}, [\theta^d, \bar{\theta}]\}$ and monotonically decreases in θ_2 to a point $(\theta^d, \bar{\theta})$, while $h(\underline{\theta}, \theta_1)$ starts from the interior of a segment $\{[\theta^d, \bar{\theta}], \bar{\theta}\}$ and monotonically decreases in θ_1 to a point $(\bar{\theta}, \theta^d)$. If the condition $\forall y < \theta^*$, $h(\underline{\theta}, h(\underline{\theta}, y)) < y$ holds, $h(\underline{\theta}, \theta_2)$ is below $h(\underline{\theta}, \theta_1)$ in the area above the 45 degree line, while $h(\underline{\theta}, \theta_2)$ is above $h(\underline{\theta}, \theta_1)$ in the area below the 45 degree line. Therefore, θ^* is the only intersection of $h(\underline{\theta}, \theta_2)$ and $h(\underline{\theta}, \theta_1)$. If the condition does not hold, $h(\underline{\theta}, \theta_2)$ and $h(\underline{\theta}, \theta_1)$ have an intersection below and above the 45 degree line and θ^* is not a unique (2,2)-rule.

(b) If $h(\underline{\theta}, y)$ is concave, both $h(\underline{\theta}, \theta_2)$ and $h(\underline{\theta}, \theta_1)$ must lie above the straight line connecting two points $(\theta^d, \bar{\theta})$ and $(\bar{\theta}, \theta^d)$ (otherwise they cannot be connecting the two points

described above). Call this line the negative 45 degree line. Because the point $\left(\frac{\theta^d + \bar{\theta}}{2}, \frac{\theta^d + \bar{\theta}}{2}\right)$ is on the negative 45 line, $\theta^* \in \left(\frac{\theta^d + \bar{\theta}}{2}, \bar{\theta}\right)$.

(iii) $h(\underline{\theta}, \theta_2)$ starts from the interior of a segment $\{\bar{\theta}, [\underline{\theta}, \theta^d]\}$ and decreases in θ_2 to a point $(\theta^d, \bar{\theta})$, while $h(\underline{\theta}, \theta_1)$ starts from the interior of a segment $\{[\underline{\theta}, \theta^d], \bar{\theta}\}$ and decreases in θ_1 to a point $(\bar{\theta}, \theta^d)$. The rest of the proof is analogous to (ii) and hence omitted.

(iv) $h(\underline{\theta}, \theta_2)$ starts from a point $(\bar{\theta}, \theta^d)$ and decreases in θ_2 to a point $(\theta^d, \bar{\theta})$, while $h(\underline{\theta}, \theta_1)$ starts from a point $(\theta^d, \bar{\theta})$ and decreases in θ_1 to a point $(\bar{\theta}, \theta^d)$. The rest of the proof is analogous to (ii) and (iii) and hence omitted.

(v) If $\{\theta \in \Theta | v_1(\theta) = 0\} \cap (\underline{\theta}, \bar{\theta})^2$ is empty, either $\{\theta \in \Theta | v_1(\theta) = 0\} = (\underline{\theta}, \underline{\theta})$ or $\{\theta \in \Theta | v_1(\theta) = 0\} = (\bar{\theta}, \bar{\theta})$. For the former, $h(\underline{\theta}, \theta_2) = \underline{\theta} \forall \theta_2 \in \Theta_2$ while for the latter $h(\theta_2, \bar{\theta}) = \bar{\theta} \forall \theta_2 \in \Theta_2$. Hence, there is no posterior implementable (2, 2)-rule. If a posterior implementable (2, 2)-rule exists, there must be $(\theta_{1,1}, \theta_{2,1}) \in (\underline{\theta}, \bar{\theta})^2$ such that either $\theta_{1,1} = h(\underline{\theta}, \theta_{2,1})$ and $\theta_{2,1} = h(\underline{\theta}, \theta_{1,1})$, or $\theta_{1,1} = h(\theta_{2,1}, \bar{\theta})$ and $\theta_{2,1} = h(\theta_{1,1}, \bar{\theta})$. For the former, $v_1(\theta) = 0$ must go through the interior of a vertical segment $\{\theta \in \Theta | \theta_1 = \theta_{1,1}, \theta_2 \in [\underline{\theta}, \theta_{2,1}]\}$. For the latter, it must go through a vertical segment $\{\theta \in \Theta | \theta_1 = \theta_{1,1}, \theta_2 \in [\theta_{2,1}, \bar{\theta}]\}$. ■(L4)

Proof of Proposition.

(i) Consider a posterior implementable (K, L) -rule with $K + L \geq 6$. At least one posterior type set $\hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l}$ in which the high outcome ϕ^+ is chosen is characterized by

$$\begin{aligned} \underline{\theta}'_1 &\equiv \min \hat{\Theta}_{1,k} = h_1(\theta_{2,l-1}, \theta_{2,l}) \text{ and } \bar{\theta}'_1 \equiv \max \hat{\Theta}_{1,k} = h_1(\theta_{2,l-2}, \theta_{2,l-1}), \\ \underline{\theta}'_2 &\equiv \min \hat{\Theta}_{2,l} = h_2(\theta_{1,k-1}, \theta_{1,k}) \text{ and } \bar{\theta}'_2 \equiv \max \hat{\Theta}_{2,l} = h_2(\theta_{1,k-2}, \theta_{1,k-1}). \end{aligned}$$

Define $h_1 : \hat{\Theta}_{2,l} \rightarrow \hat{\Theta}_{1,k}$ by $h_1(\underline{\theta}'_2, \theta_2)$ for $\theta_2 \in \hat{\Theta}_{2,l}$. Similarly define $h_2 : \hat{\Theta}_{1,k} \rightarrow \hat{\Theta}_{2,l}$ by $h_2(\theta'_1, \theta_1)$ for $\theta_1 \in \hat{\Theta}_{1,k}$. Note that h_1 and h_2 are continuous, strictly decreasing in the interior of $\hat{\Theta}_{k,l}$, $h_1(\underline{\theta}'_2, \bar{\theta}'_2) = h_1(\theta_{2,l-1}, \theta_{2,l}) = \underline{\theta}'_1$ and $h_2(\theta'_1, \bar{\theta}'_1) = h_2(\theta_{1,k-1}, \theta_{1,k}) = \underline{\theta}'_2$. Also, $h_1(\underline{\theta}'_2, \underline{\theta}'_2) = \{\theta_1 \in \Theta_1 | v_1(\theta_1, \underline{\theta}'_2) = 0\} \in (\underline{\theta}'_1, \bar{\theta}'_1)$, i.e., $v_1(\theta) = 0$ goes through a horizontal segment $\{\theta | \theta_1 \in \hat{\Theta}_{1,k}, \theta_2 = \underline{\theta}'_2\}$, because $v_1(\theta) = 0$ goes through two vertical segments $\{\theta | \theta_1 = \underline{\theta}'_1, \theta_2 \in \hat{\Theta}_{2,l}\}$ and $\{\theta | \theta_1 = \bar{\theta}'_1, \theta_2 \in \hat{\Theta}_{2,l-1}\}$ (otherwise $\underline{\theta}'_1$ and $\bar{\theta}'_1$ would not be indifferent types). Similarly, $h_2(\underline{\theta}'_1) = \{\theta_2 \in \Theta_2 | v_2(\underline{\theta}'_1, \theta_2) = 0\} \in (\underline{\theta}'_2, \bar{\theta}'_2)$, i.e., $v_2(\theta) = 0$ goes through a vertical segment $\{\theta | \theta_1 = \underline{\theta}'_1, \theta_2 \in \hat{\Theta}_{2,l}\}$ for the same reason. Hence, a mapping $t : \hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l} \rightarrow \hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l}$ defined by $t(\theta_1, \theta_2) = (h_1(\underline{\theta}'_2, \theta_2), h_2(\theta'_1, \theta_1))$ has at least one fixed point in the interior of $\hat{\Theta}_{k,l}$. Let $x^* = (x_1^*, x_2^*) \in \hat{\Theta}_{1,k} \times \hat{\Theta}_{2,l}$ be such a fixed point. This is a posterior implementable (2, 2)-rule defined on $\hat{\Theta}_{1,k} \times \hat{\Theta}_{2,k}$. Therefore, a (K, L) -rule with $K + L \geq 6$ is not PRP.

(ii) (If) If (a) holds, by **Lemma 2(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, by **Lemma 2(ii)**, there is no posterior implementable rule that is not constant. Also, unanimous improvement by a constant rule is impossible because $\phi^0 = 1$. 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 are posterior implementable, 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 posterior disagreement on the set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$ and there is posterior agreement on d_1 on the set $[\theta_1^d, \bar{\theta}_1] \times \Theta_2$. By **Proposition (ii)** applied to each set, 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 the set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$ and interim agreement on d_1 in the set $[\bar{\theta}_1, \theta_1^d] \times \Theta_2$. There is no improvement by a constant rule because $\phi^- = 0$ on $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$ and $\phi^+ = 1$ on the set $[\bar{\theta}_1, \theta_1^d] \times \Theta_2$. By **Lemma 2** applied to each set, there is no posterior implementable low type (2, 2)-rule in the set $[\underline{\theta}_1, \theta_1^d] \times \Theta_2$ and there is no posterior implementable high type (2, 2)-rule in the set $[\bar{\theta}_1, \theta_1^d] \times \Theta_2$. Thus, the last two statements in (c) imply that there is no more information revelation in each case.

(Only if) If none of (a)-(c) holds, either (a') $h_2(\underline{\theta}_1, \theta_1^d) \in (\underline{\theta}_2, \bar{\theta}_2)$ or $h_2(\theta_1^d, \bar{\theta}_1) \in (\underline{\theta}_2, \bar{\theta}_2)$, (b') ($h_2(\underline{\theta}_1, \theta_1^d) = \underline{\theta}_2$ and $\phi^+ < 1$) or ($h_2(\theta_1^d, \bar{\theta}_1) = \bar{\theta}_2$ and $\phi^- > 0$), (c') $h_2(\underline{\theta}_1, \theta_1^d) = \bar{\theta}_2$, $h_2(\theta_1^d, \bar{\theta}_1) = \underline{\theta}_2$ and ($\phi^+ < 1$ or $\phi^- > 0$ or a high type (2, 2)-rule is posterior implementable in the set $[\underline{\theta}_1, \theta_1^d]$ or a low type (2, 2)-rule is posterior implementable in the set $[\theta_1^d, \bar{\theta}_1]$). None of these cases satisfies the definition of PRP.

(iv) (If) In the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$, no (2, 1)-rule is posterior implementable because $\theta_{1,1} = \min \hat{\Theta}_{1,2}$ is an indifferent type of agent 1 believing that agent 2's type is in the set $\hat{\Theta}_{2,1}$. Similarly, in the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,2}$, no (1, 2)-rule is posterior implementable because $\theta_{2,1} = \min \hat{\Theta}_{2,2}$ is an indifferent type of agent 2 believing that agent 1's type is in the set $\hat{\Theta}_{1,1}$. In the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$, there is interim agreement on d_0 . From the proof of **Lemma 2**, there is no high type (2, 2)-rule in the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$. Therefore, (a)-(c) is sufficient for no information revelation. Because there is posterior agreement on d_0 in the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$ and posterior agreement on d_1 in the other three sets, $\phi^- = 0$ and $\phi^+ = 1$ imply that there is no improvement by constant rules.

(Only if) Necessity of (a)-(c) is obvious from the definition of PRP. If the condition ($\phi^- = 0$ and $\phi^+ = 1$) are violated when (a)-(c) are satisfied, improvement by constant rules is possible either in the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$ (by lower outcome than $\phi^- > 0$) or in the other sets (by higher outcome than $\phi^+ < 1$).

(v) (a) (If) A step function which represents a posterior implementable (3, 2)-rule has two vertical segments and one horizontal segment, and $v_1(\theta) = 0$ crosses the two vertical segments while $v_2(\theta) = 0$ crosses the horizontal segment. Because neither $v_1(\theta) = 0$ nor $v_2(\theta) = 0$ goes through $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$ and $\hat{\Theta}_{1,3} \times \hat{\Theta}_{2,2}$, only constant rules are posterior implementable in these two sets. In the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,2}$, there is posterior agreement on d_1 , and (2, 1)-rules, (1, 2)-rules, and high type (2, 2)-rules are not posterior implementable. Similarly, in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$, there is posterior agreement on d_0 , and (2, 1)-, (1, 2)-, and high

type (2,2)-rules are not posterior implementable. In the set $\hat{\Theta}_{1,3} \times \hat{\Theta}_{2,1}$, (2,1)- and high type (2,2)-rules are not posterior implementable. Similarly, (2,1)- and low type (2,2)-rules are not posterior implementable in the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,2}$. Hence, it suffices to show: if no low type (2,2)-rule is posterior implementable in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,2}$ and no high type (2,2)-rule is posterior implementable in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$, then no (1,2)- and low type (2,2)-rule is posterior implementable in the set $\hat{\Theta}_{1,3} \times \hat{\Theta}_{2,1}$ and no (1,2)- and high type (2,2)-rule is implementable in the set $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,2}$. Notice that if $v_2(\theta) = 0$ crosses the left vertical segment $\{\theta_{1,1}, (\theta_{2,1}, \bar{\theta}_2)\}$, then a low type (2,2)-rule is posterior implementable in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,2}$. To see this, consider $h_1(\theta_{2,1}, y_2)$ defined for $y_2 \in [\theta_{2,1}, \bar{\theta}_2]$, which connects $h_1(\theta_{2,1}, \bar{\theta}_2) = \theta_{1,1}$ and $h_1(\theta_{2,1}, \theta_{2,1}) = \left\{ \theta_1 \in \hat{\Theta}_{1,2} \mid v_1(\theta_1, \theta_{2,1}) = 0 \right\} \in (\theta_{1,1}, \theta_{1,2})$. Also consider $h_2(\theta_{1,1}, y_1)$ defined for $y_1 \in [\theta_{1,1}, \theta_{1,2}]$, which connects $h_2(\theta_{1,1}, \theta_{1,2}) = \theta_{2,1}$ and $h_2(\theta_{1,1}, \theta_{1,1}) = \left\{ \theta_2 \in \hat{\Theta}_{2,2} \mid v_2(\theta_{1,1}, \theta_2) = 0 \right\} \in (\theta_{2,1}, \bar{\theta}_2)$. They must cross at least once in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,2}$. Similarly, if $v_2(\theta) = 0$ crosses the right vertical segment $\{\theta_{1,2}, (\theta_{2,1}, \theta_{2,1})\}$, then there exists a high type (2,2)-rule in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$. Therefore, if no low type (2,2)-rule is posterior implementable in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,2}$ and no high type (2,2)-rule is posterior implementable in the set $\hat{\Theta}_{1,2} \times \hat{\Theta}_{2,1}$, then $v_2(\theta) = 0$ crosses none of the two vertical segments. This implies that $v_2(\theta) = 0$ crosses neither $\hat{\Theta}_{1,3} \times \hat{\Theta}_{2,1}$ nor $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,2}$. Therefore, no (1,2)- and (2,2)-rule is posterior implementable in these type sets. Because there is interim agreement in each posterior type set, $\phi^- = 0$ and $\phi^+ = 1$ imply there is no improvement by constant rules.

(Only if) Necessity of (2,2)-rules follows from the definition of PRP. That either $\phi^- > 0$ or $\phi^+ < 1$ clearly violates the definition of PRP because there is an improvement by constant rules either on $\hat{\Theta}_{1,3} \times \hat{\Theta}_{2,2}$ (by higher outcome than ϕ^+) or on $\hat{\Theta}_{1,1} \times \hat{\Theta}_{2,1}$ (by lower outcome than ϕ^-).

(b) The condition says $v_2(\theta) = 0$ crosses none of the two vertical segments. The necessity of this condition was proved in (a). ■(P)

Proof of Corollary.

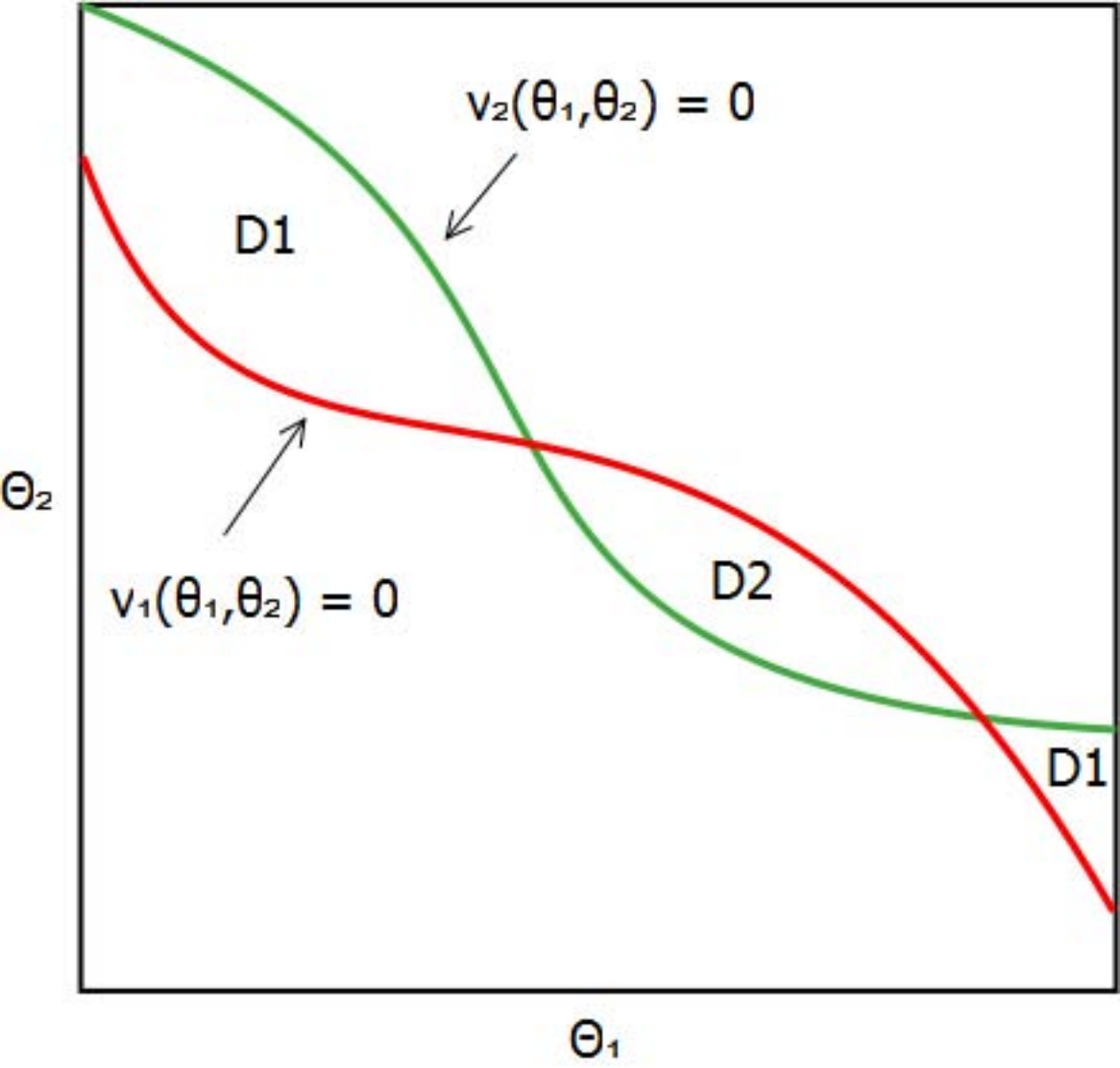
(i) With symmetric agents, **Proposition(ii)(a)** cannot occur. By **Lemma 4(i)(v)**, no information revelation is possible if and only if $\{\theta \in \Theta \mid v_1(\theta) = 0\} \cap (\underline{\theta}, \bar{\theta})^2$ is empty. There are two cases to consider. If $v_1(\theta) \geq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]^2$, then there is interim agreement on d_1 . Hence, $\phi^0 = 1$ implies PRP and $\phi^0 < 1$ implies not PRP. If $v_1(\theta) \leq 0 \forall \theta \in [\underline{\theta}, \bar{\theta}]$, then there is interim agreement on d_0 . Hence, $\phi^0 = 0$ implies PRP and $\phi^0 > 0$ implies not PRP.

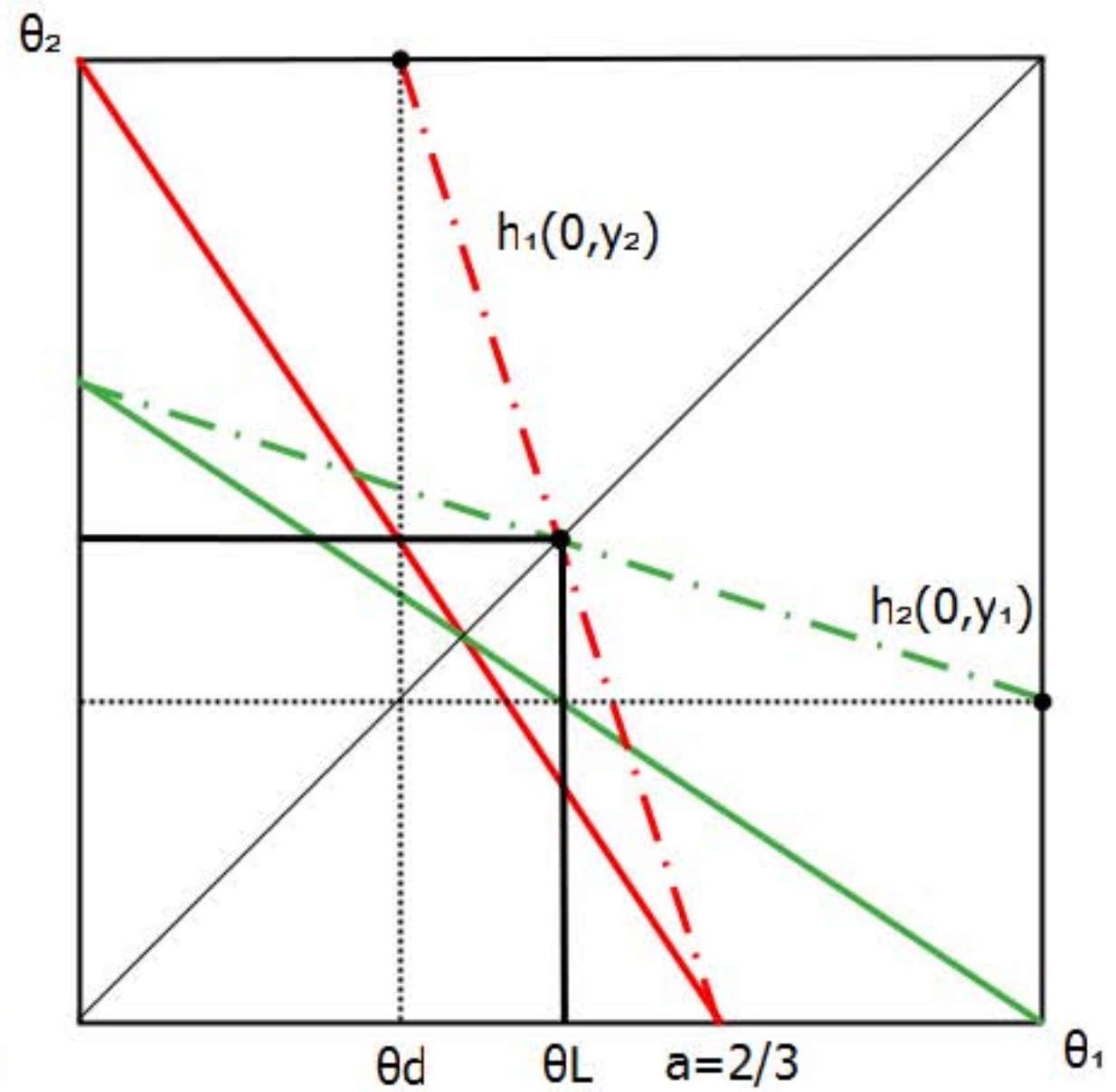
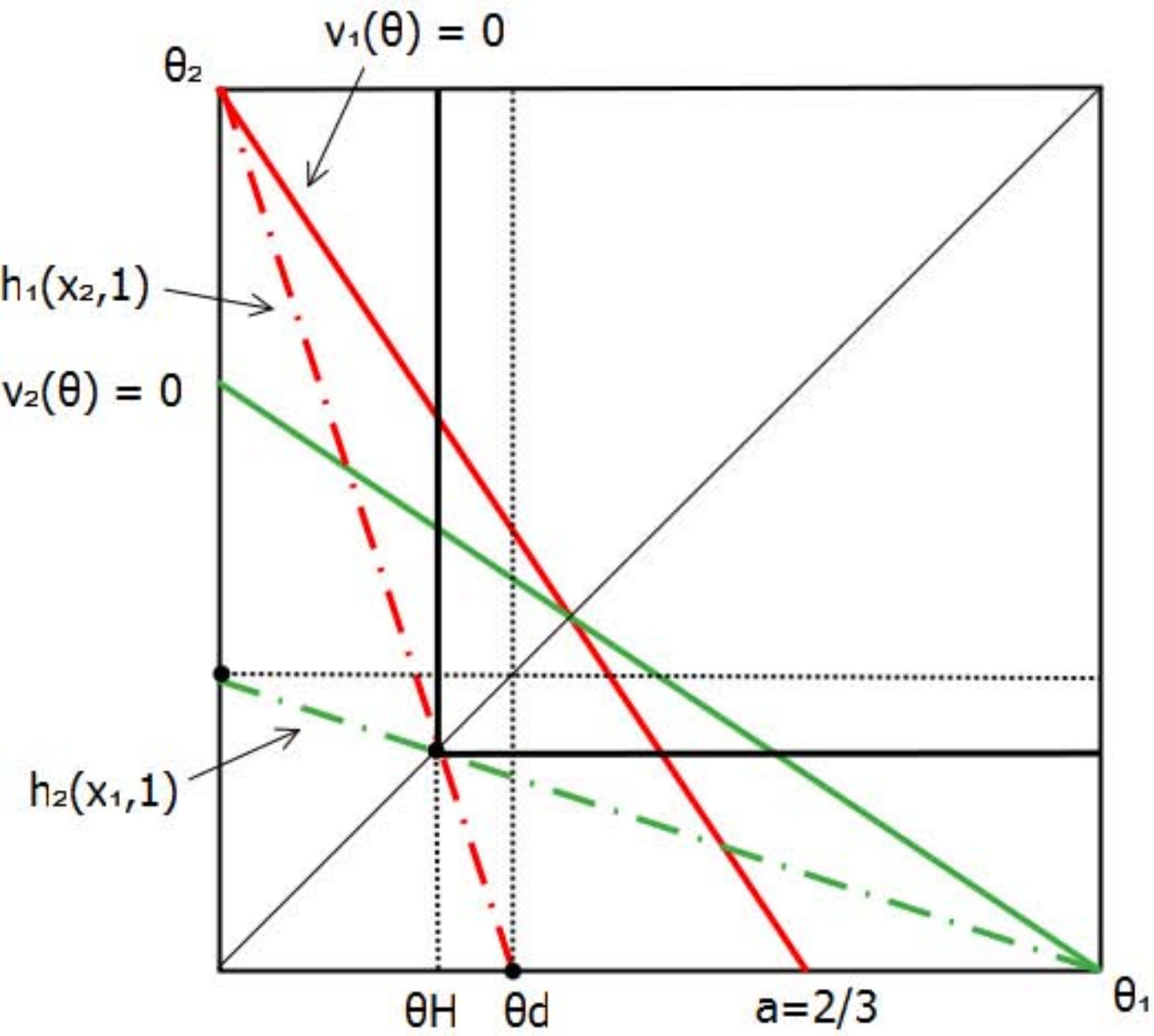
(ii) With symmetric agents, **Proposition(iii)(a)(b)** cannot occur.

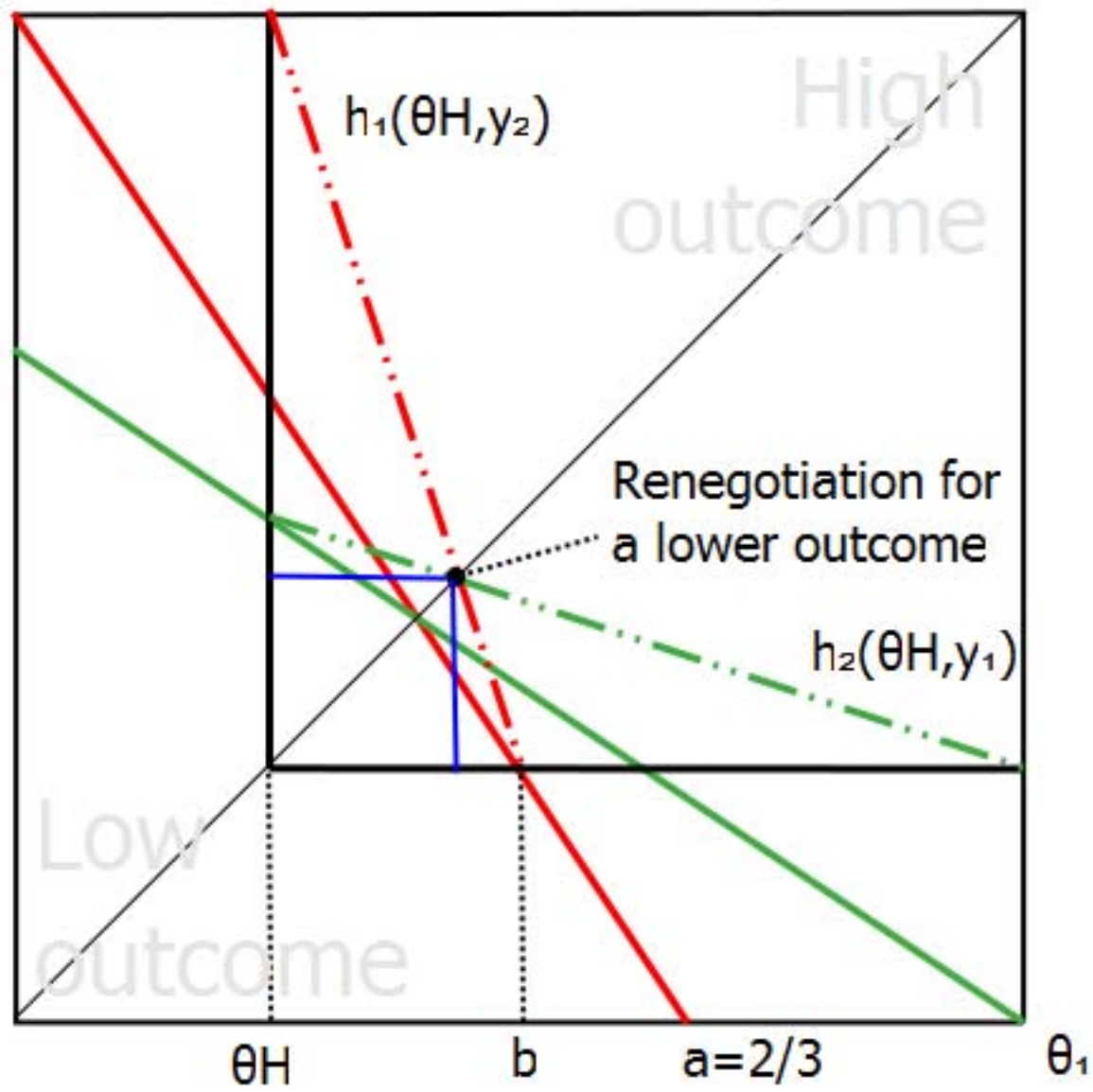
(iii) If $\theta_{1,1} = \theta_{2,1}$, two partitions $[\underline{\theta}, \theta_{1,1}] \times [\underline{\theta}, \theta_{2,1}]$ and $[\theta_{1,1}, \bar{\theta}] \times [\theta_{2,1}, \bar{\theta}]$ are both symmetric and agents' beliefs remain symmetric. If the (2,2)-rule is of low type, $v_1(\theta) = 0$ goes through the interior of $[\underline{\theta}, \theta_{1,1}] \times [\underline{\theta}, \theta_{2,1}]$, and by **Lemma 4(i)(v)** a (2,2)-rule is posterior implementable in this set. If the (2,2)-rule is of high type, $v_1(\theta) = 0$ goes through the interior of $[\theta_{1,1}, \bar{\theta}] \times [\theta_{2,1}, \bar{\theta}]$, and by **Lemma 4(i)(v)** a (2,2)-rule is posterior implementable in this set. ■(C)

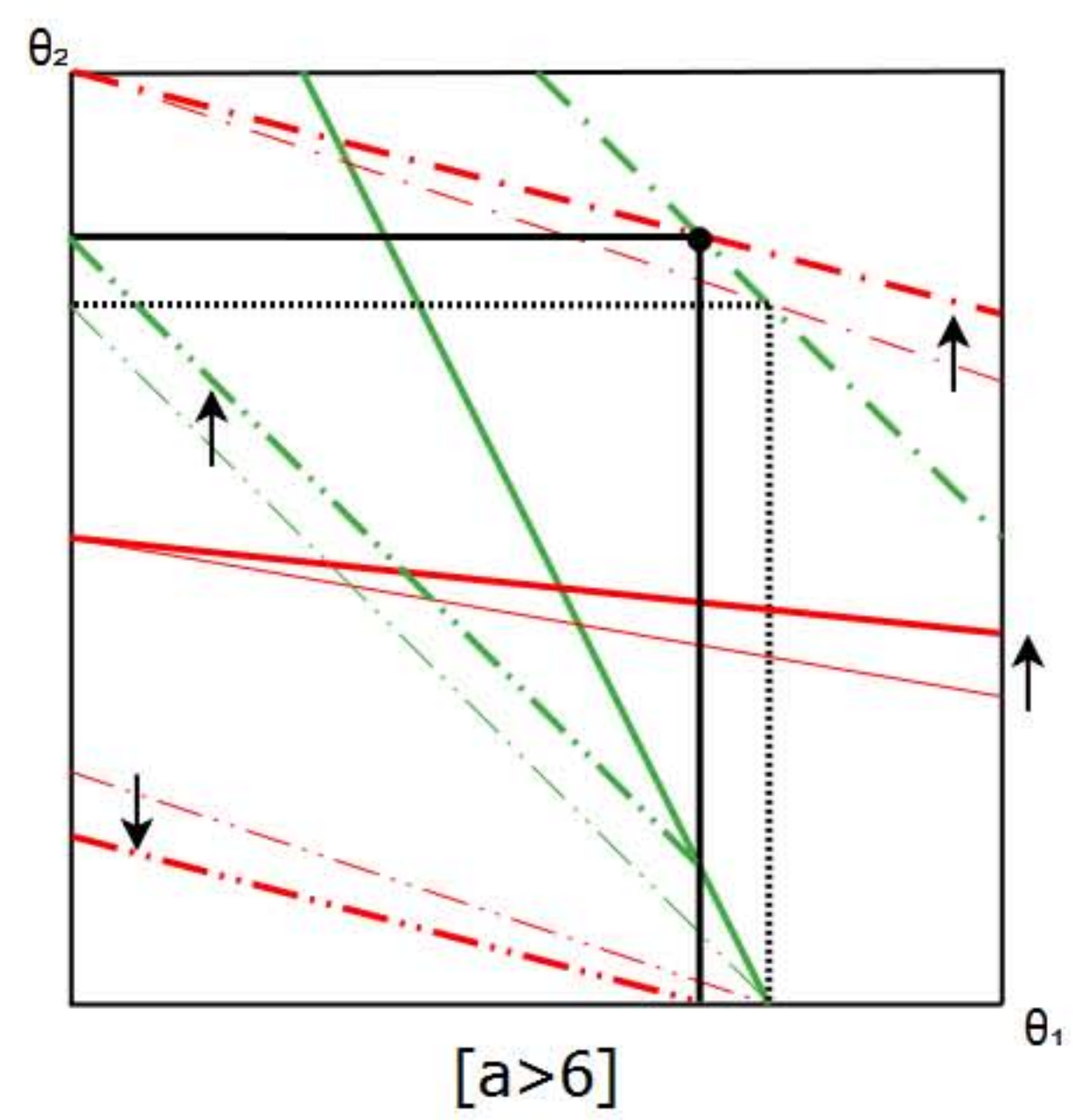
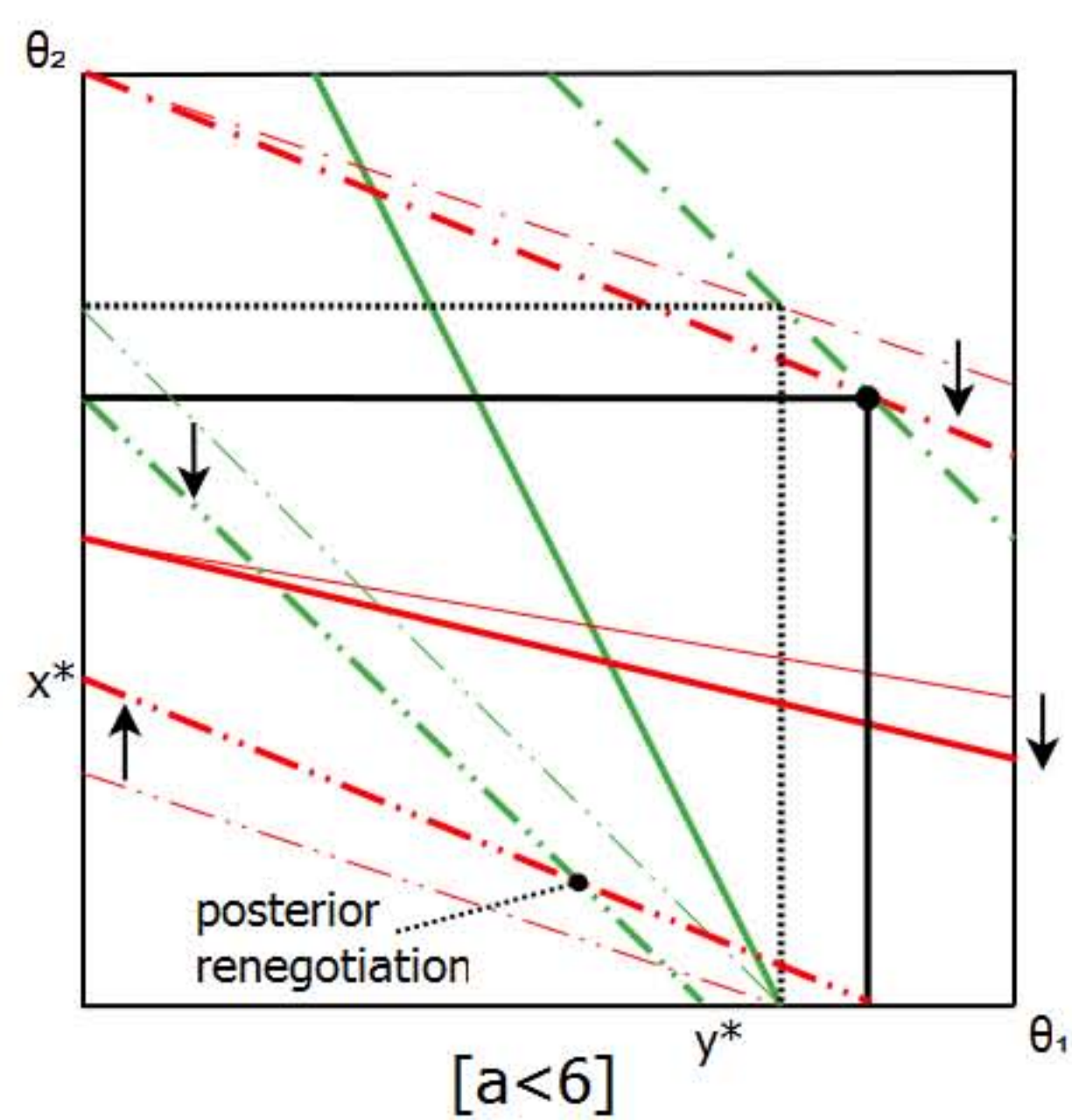
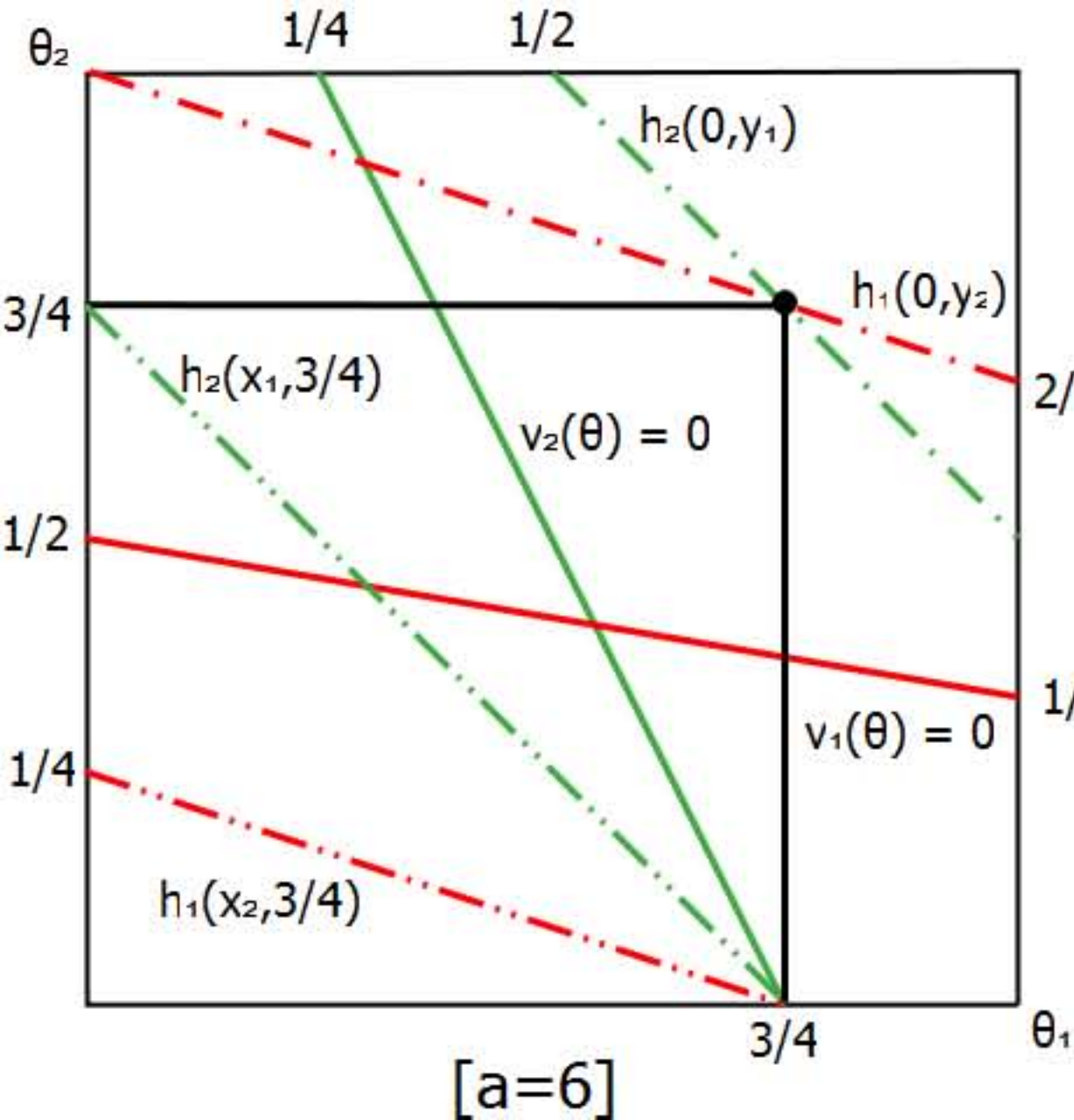
References

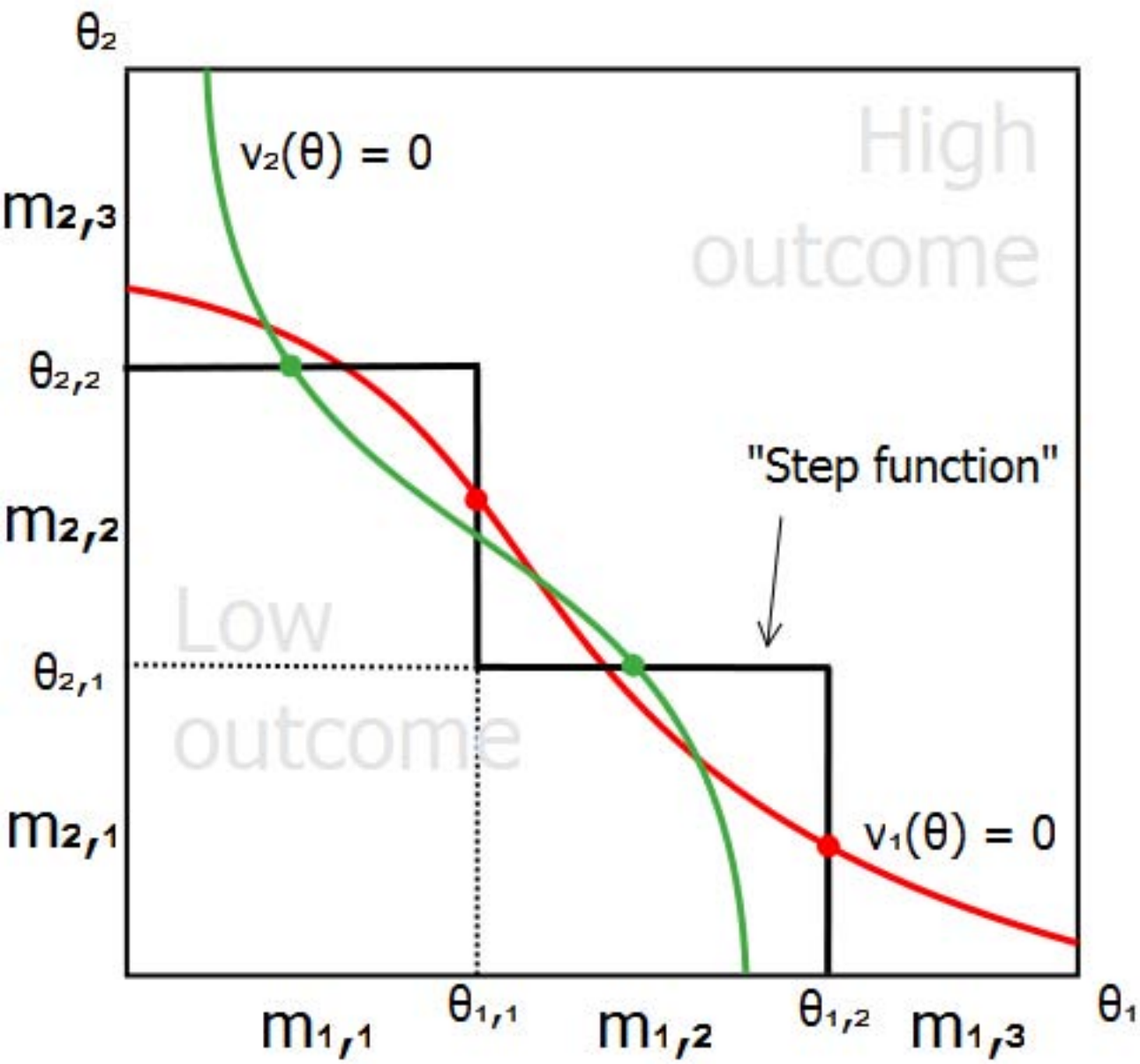
- [1] Cramton, P.C. and T.R. Palfrey (1995): “Ratifiable Mechanisms: Learning from Disagreement,” *Games and Economic Behavior*, 10, 255-283.
- [2] Forges, F. (1994): “Posterior Efficiency,” *Games and Economic Behavior*, 6, 238-261.
- [3] Green, J. R. and J.J. Laffont (1987): “Posterior Implementability 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] Poitevin, M. (2000): “Can the Theory of Incentives Explain Decentralization?,” *Canadian Journal of Economics*, 33, 4, 878-906.
- [9] Vartiainen, H. (2013): “Auction Design without Commitment,” *Journal of the European Economic Association*, 11, 2, 316-342.
- [10] Watson, J. (1999): “Starting Small and Renegotiation,” *Journal of Economic Theory*, 85, 52-90.

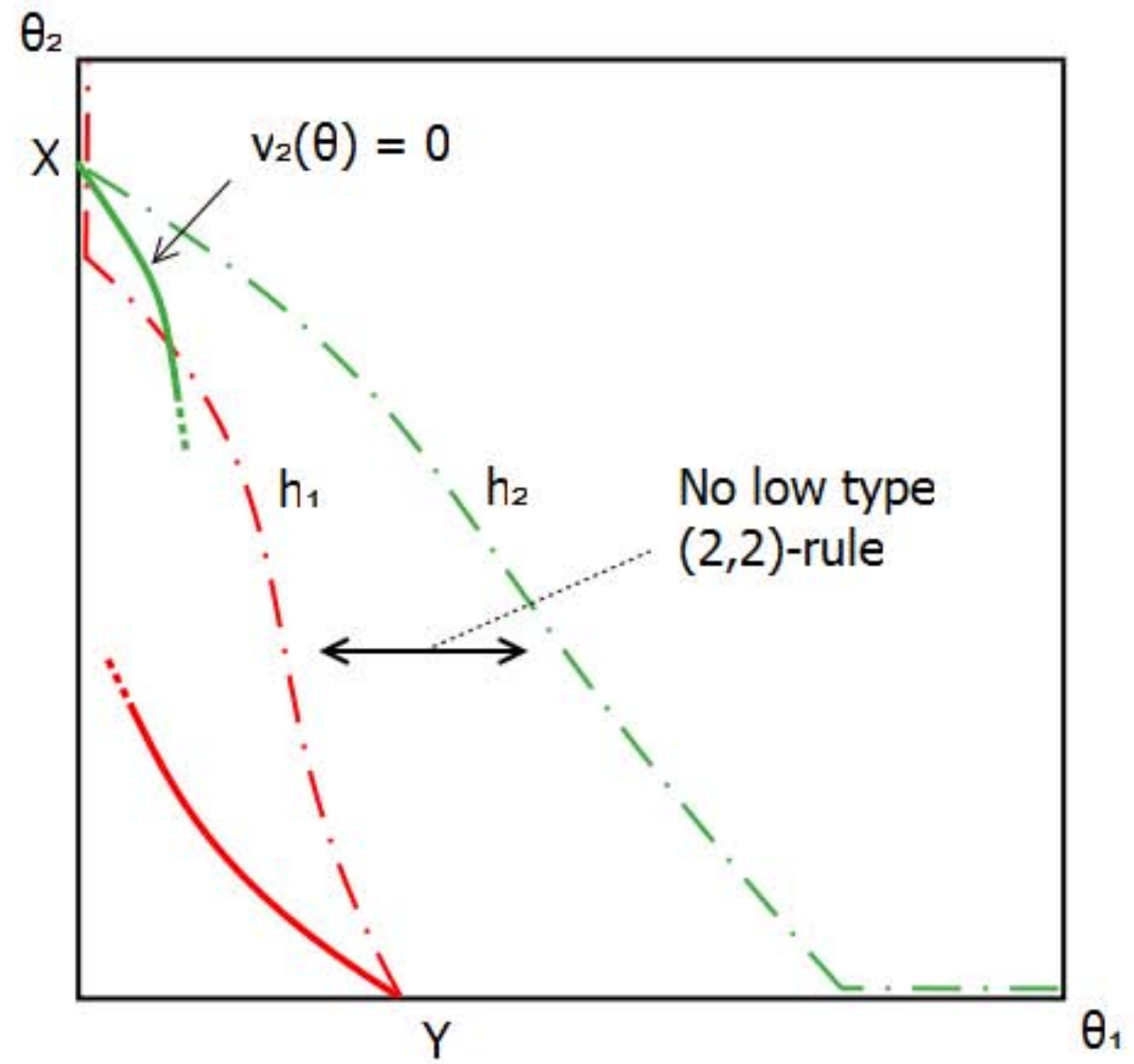
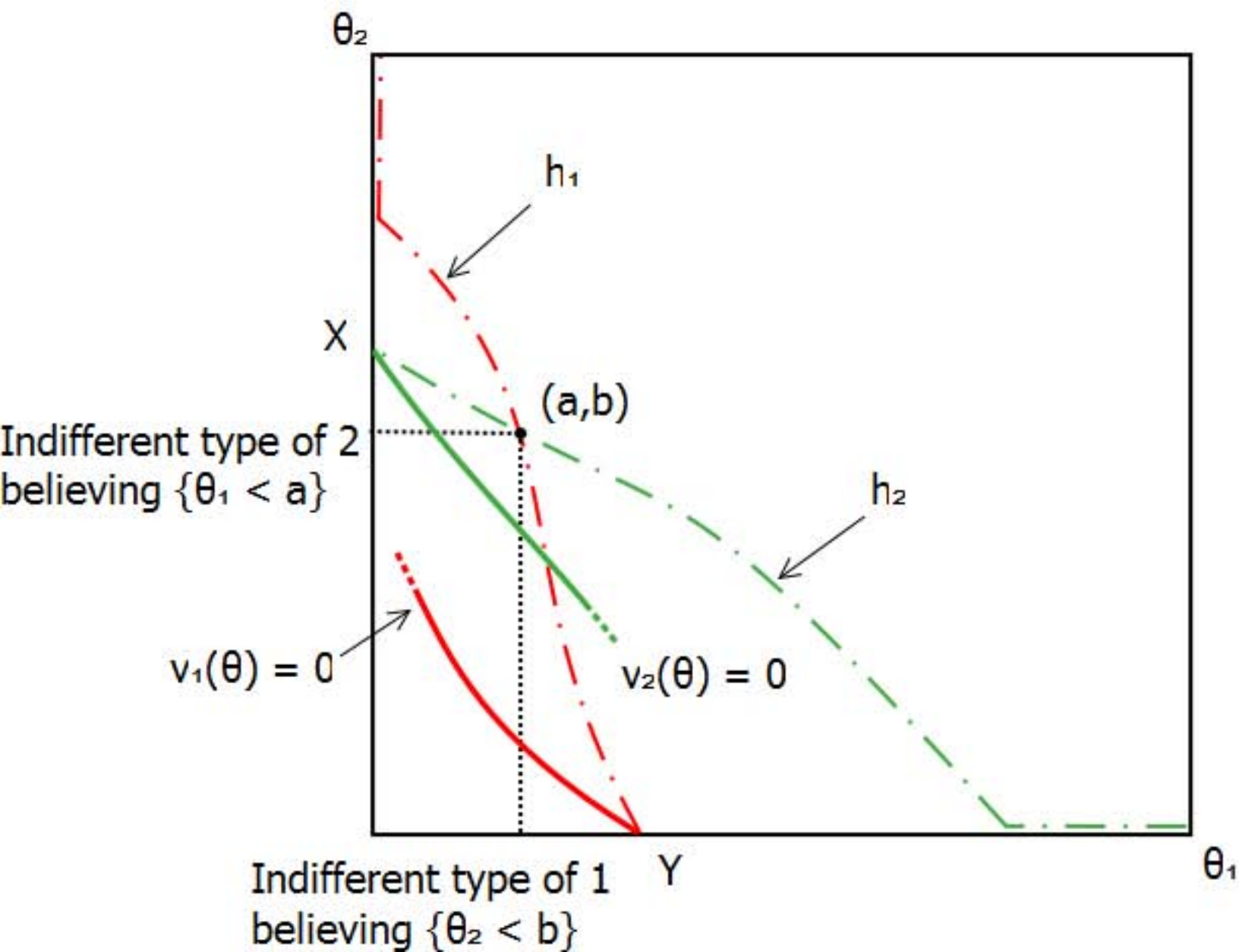


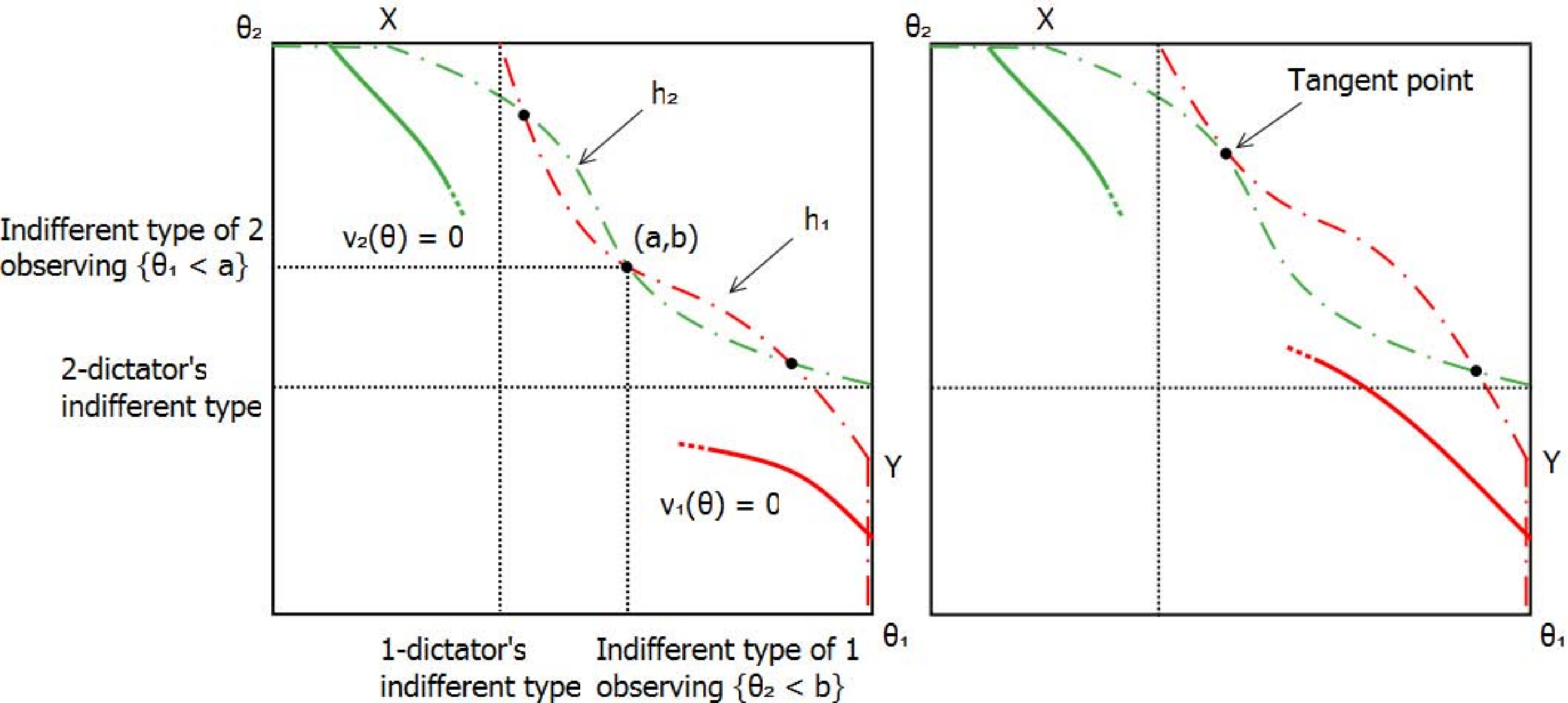


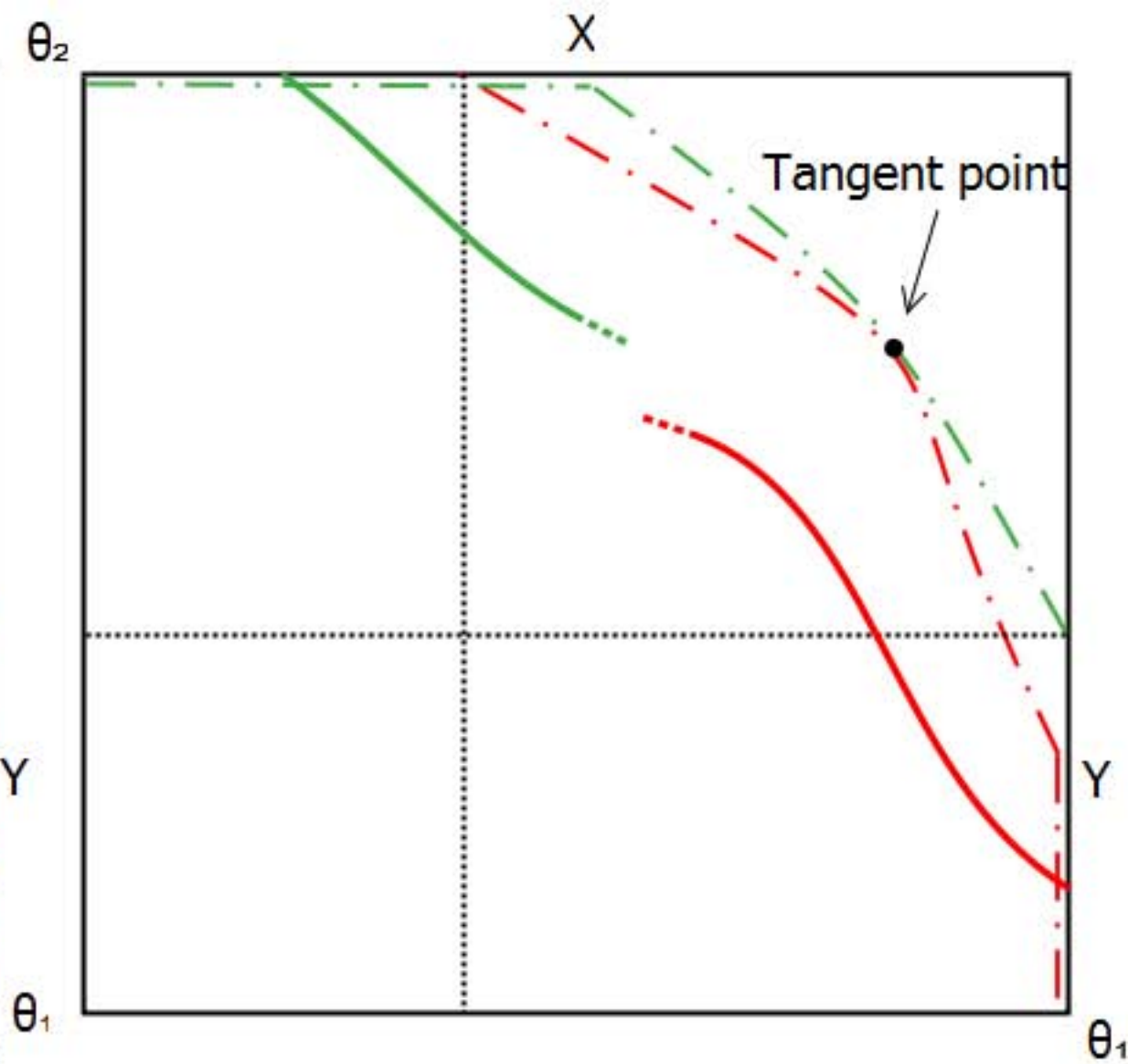
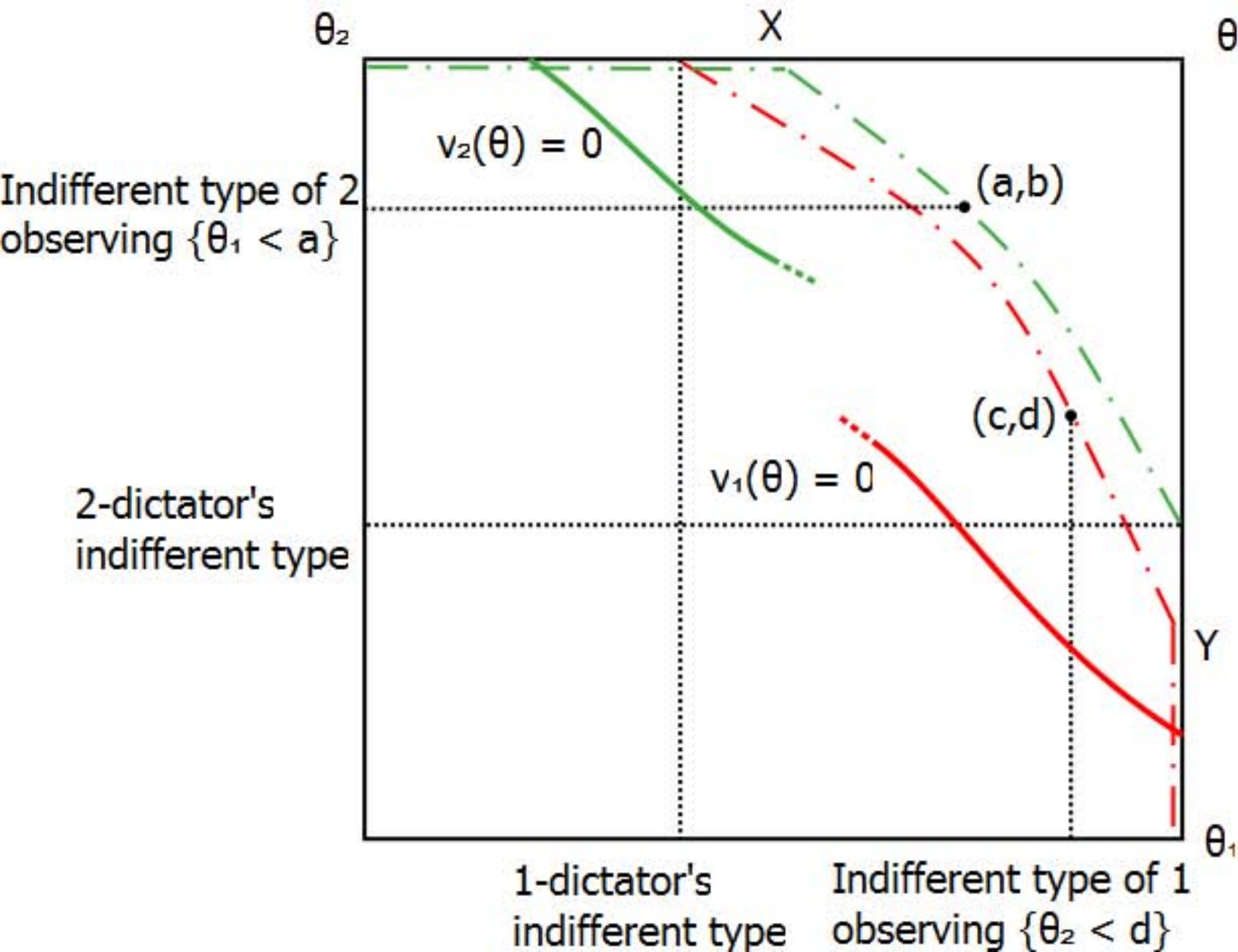
θ_2 











θ_2

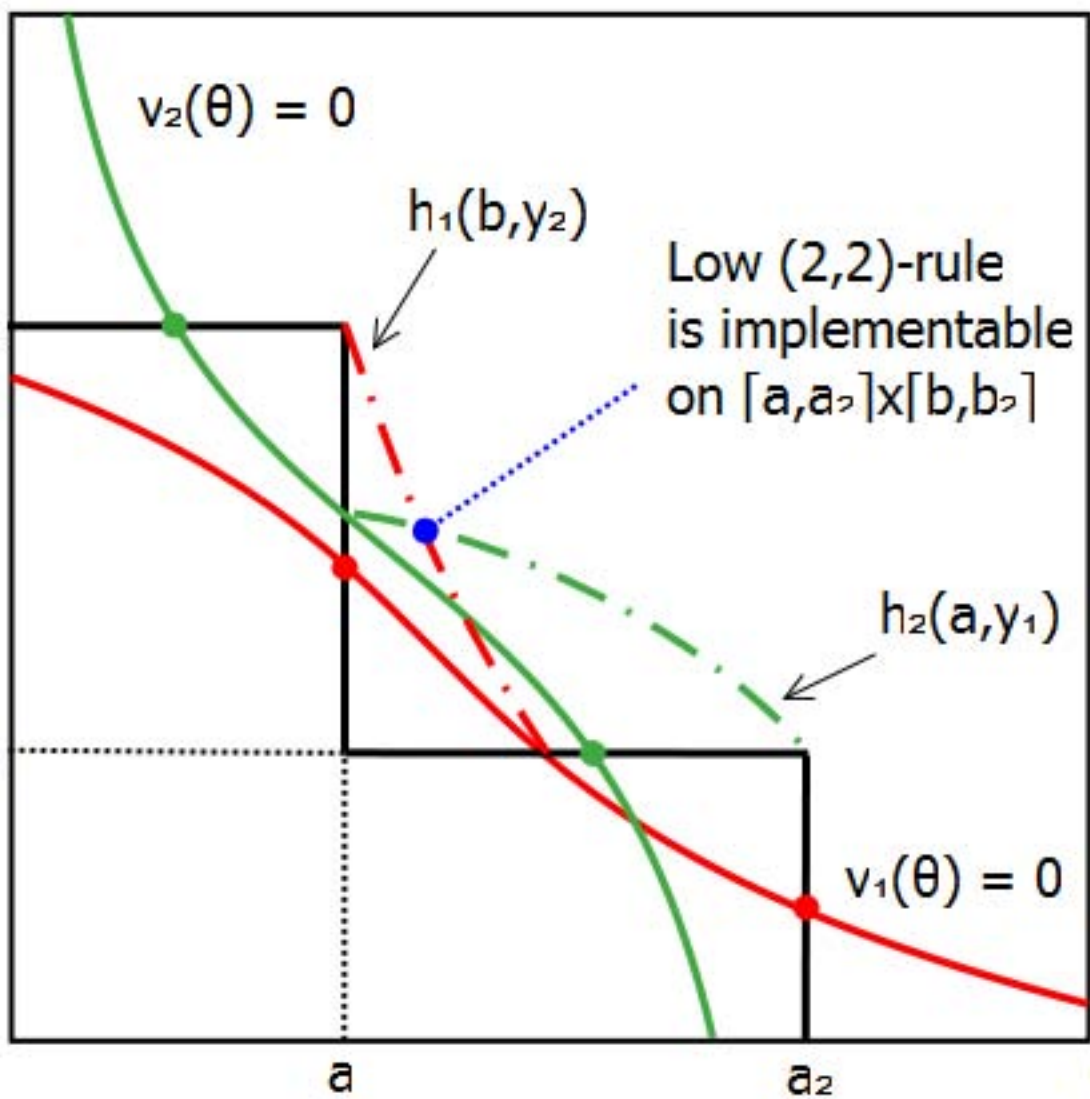
$$v_2(\theta) = 0$$

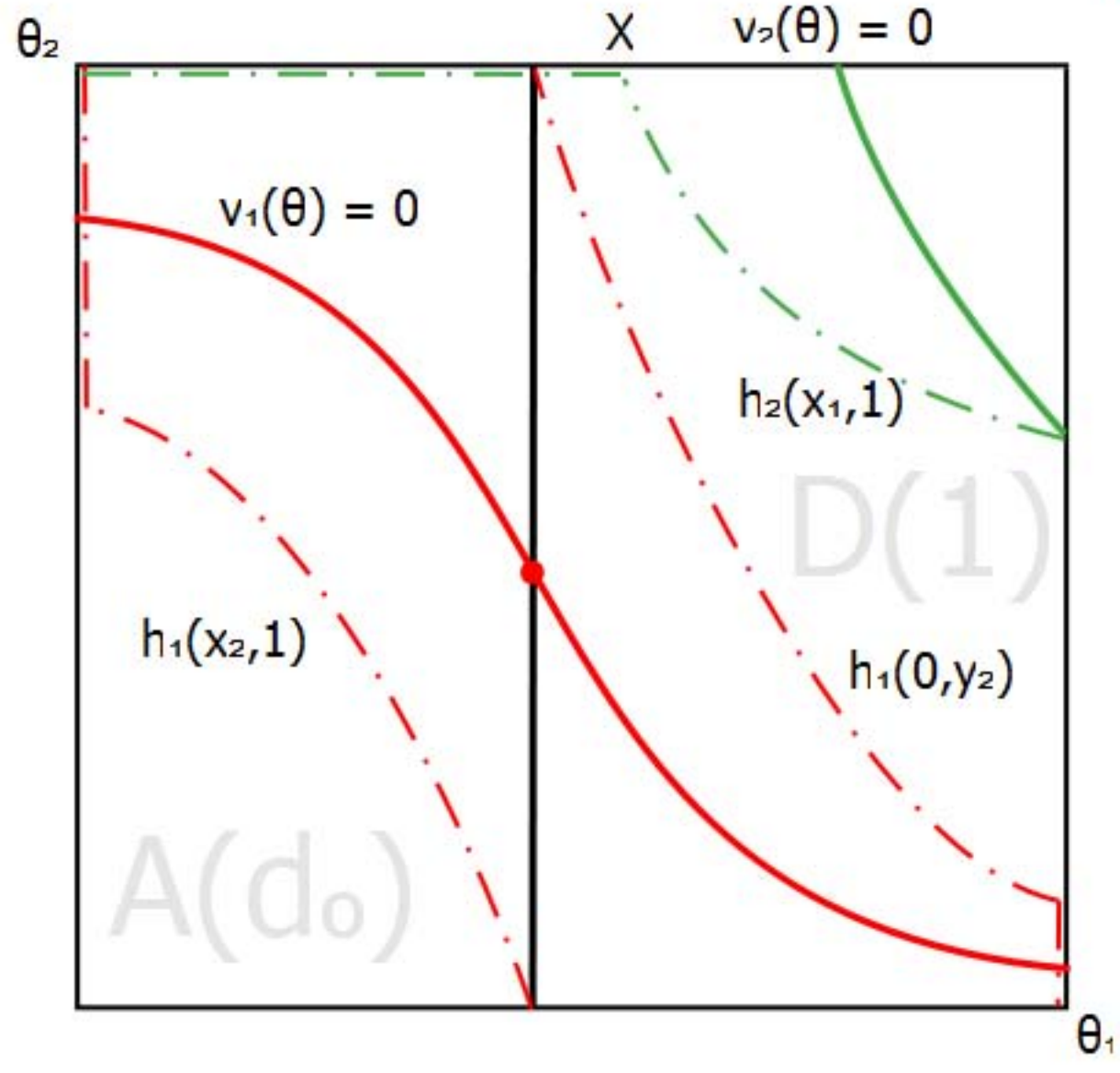
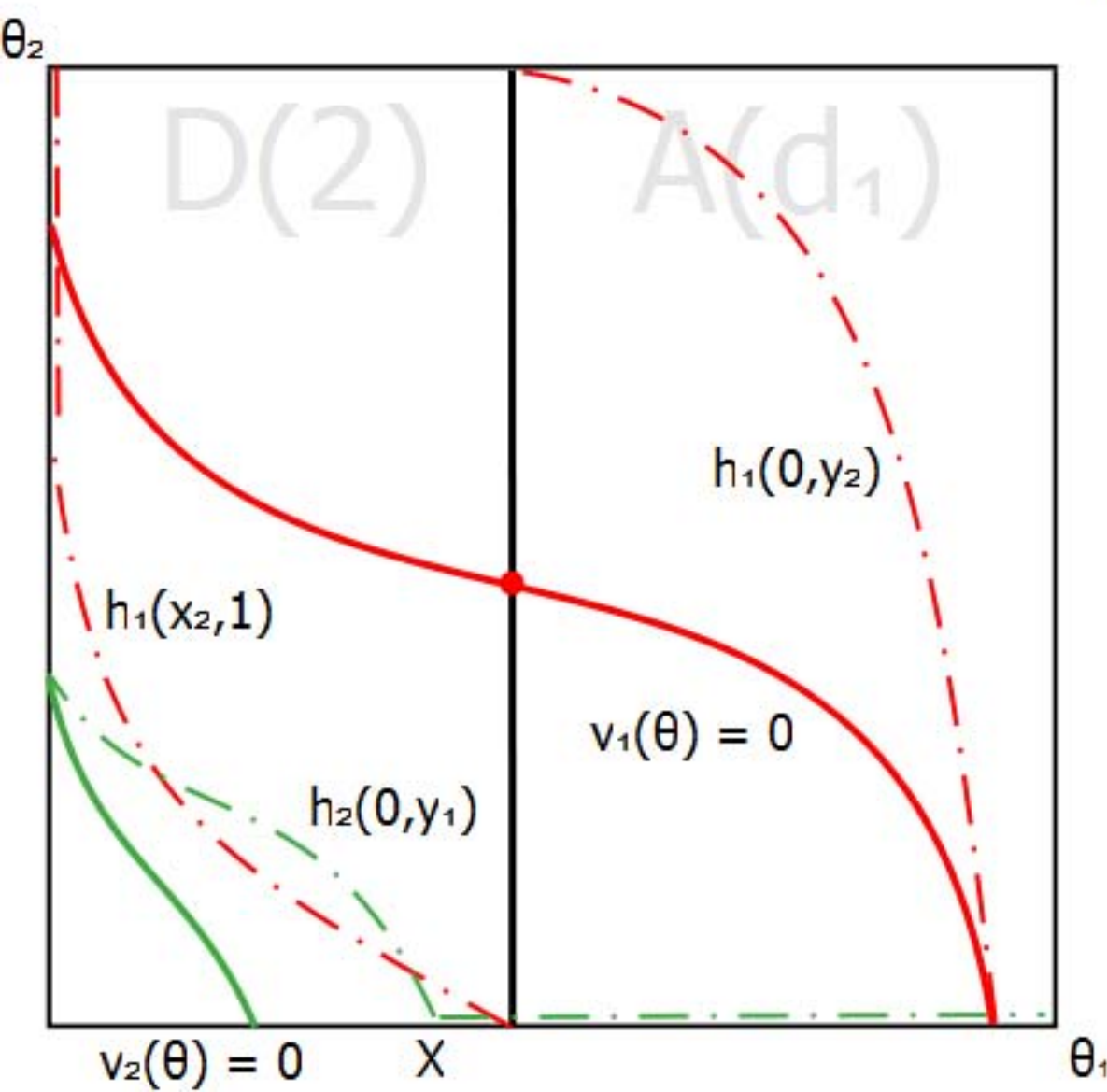
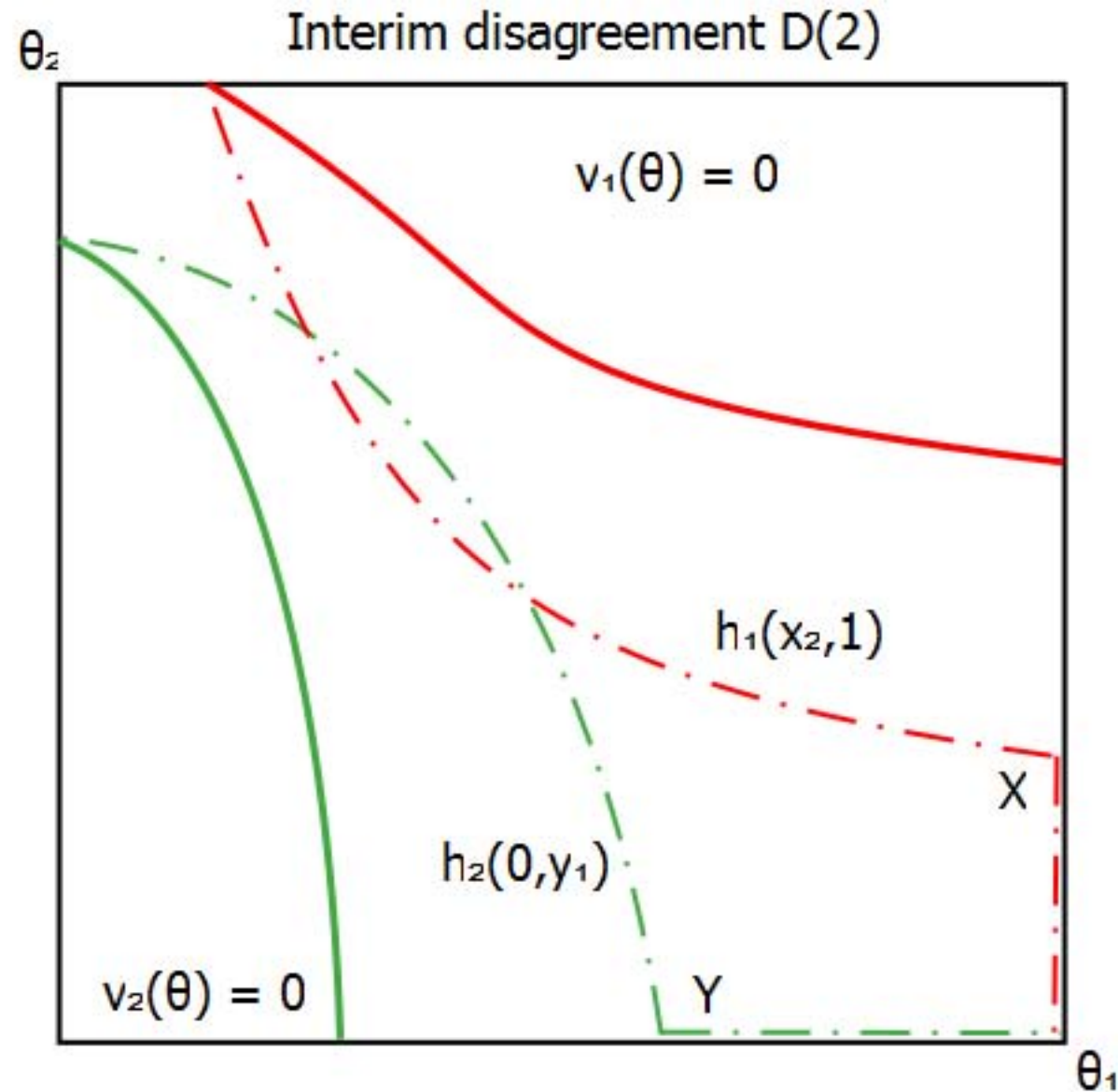
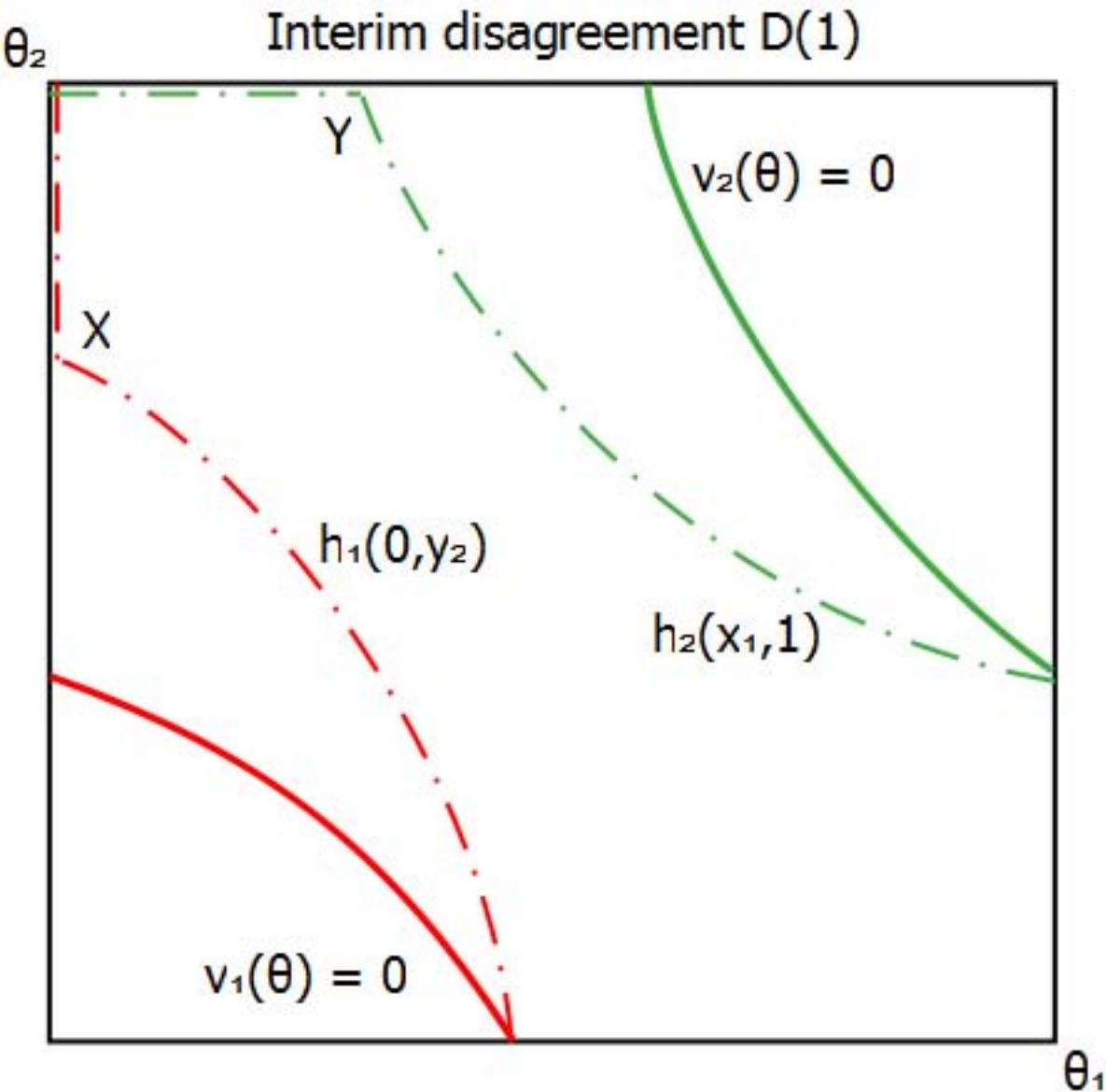
$$h_1(b, y_2)$$

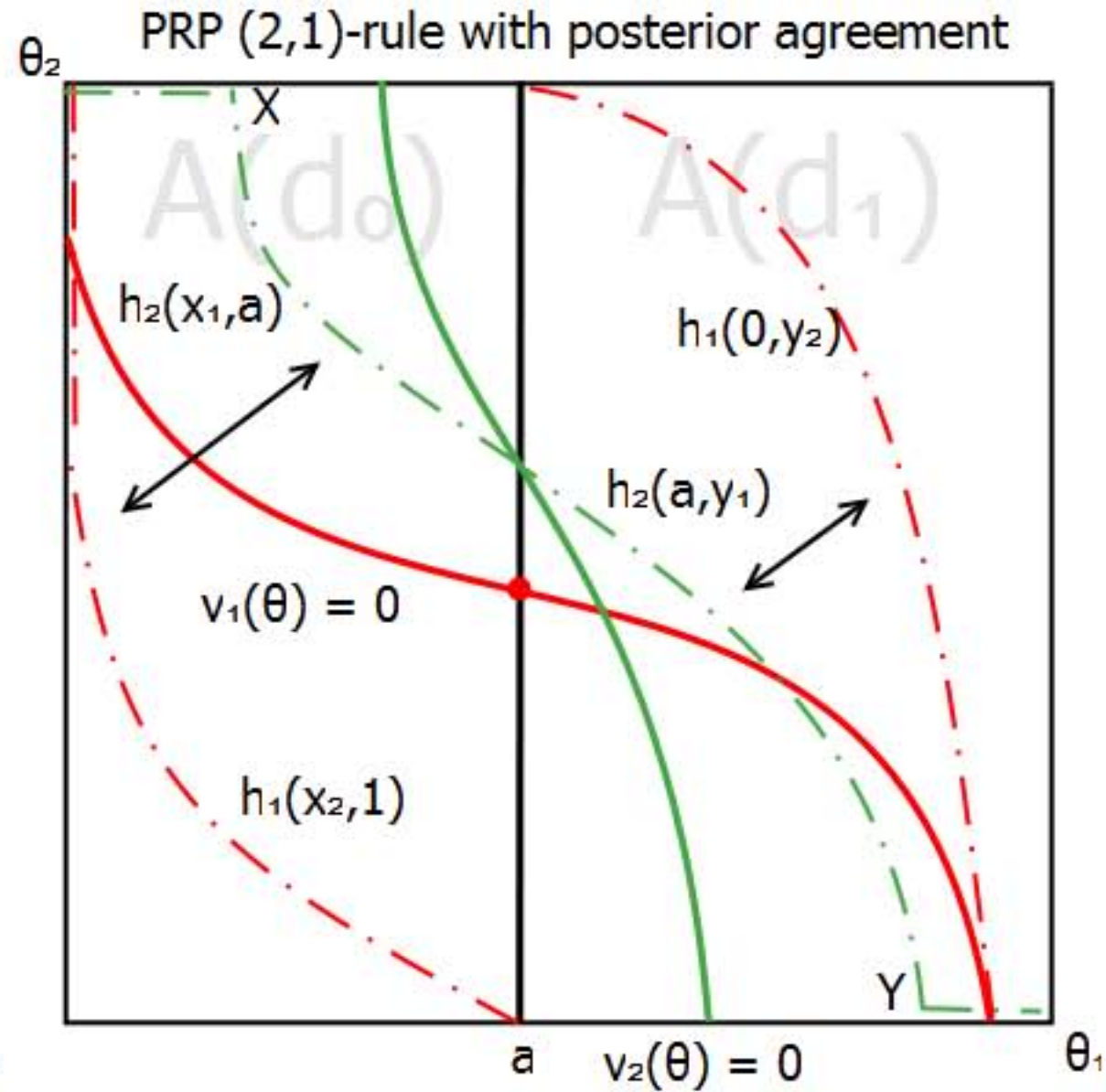
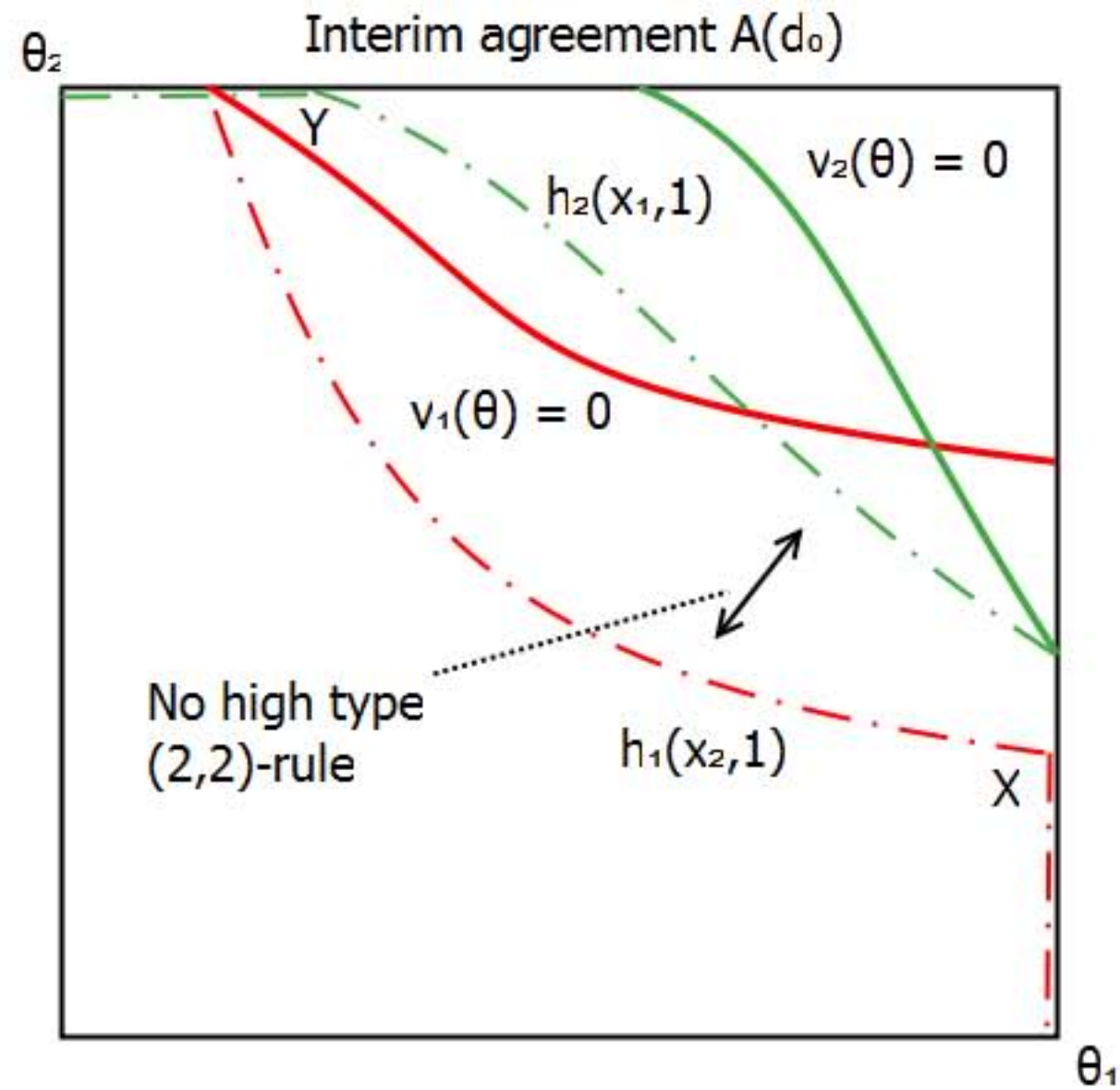
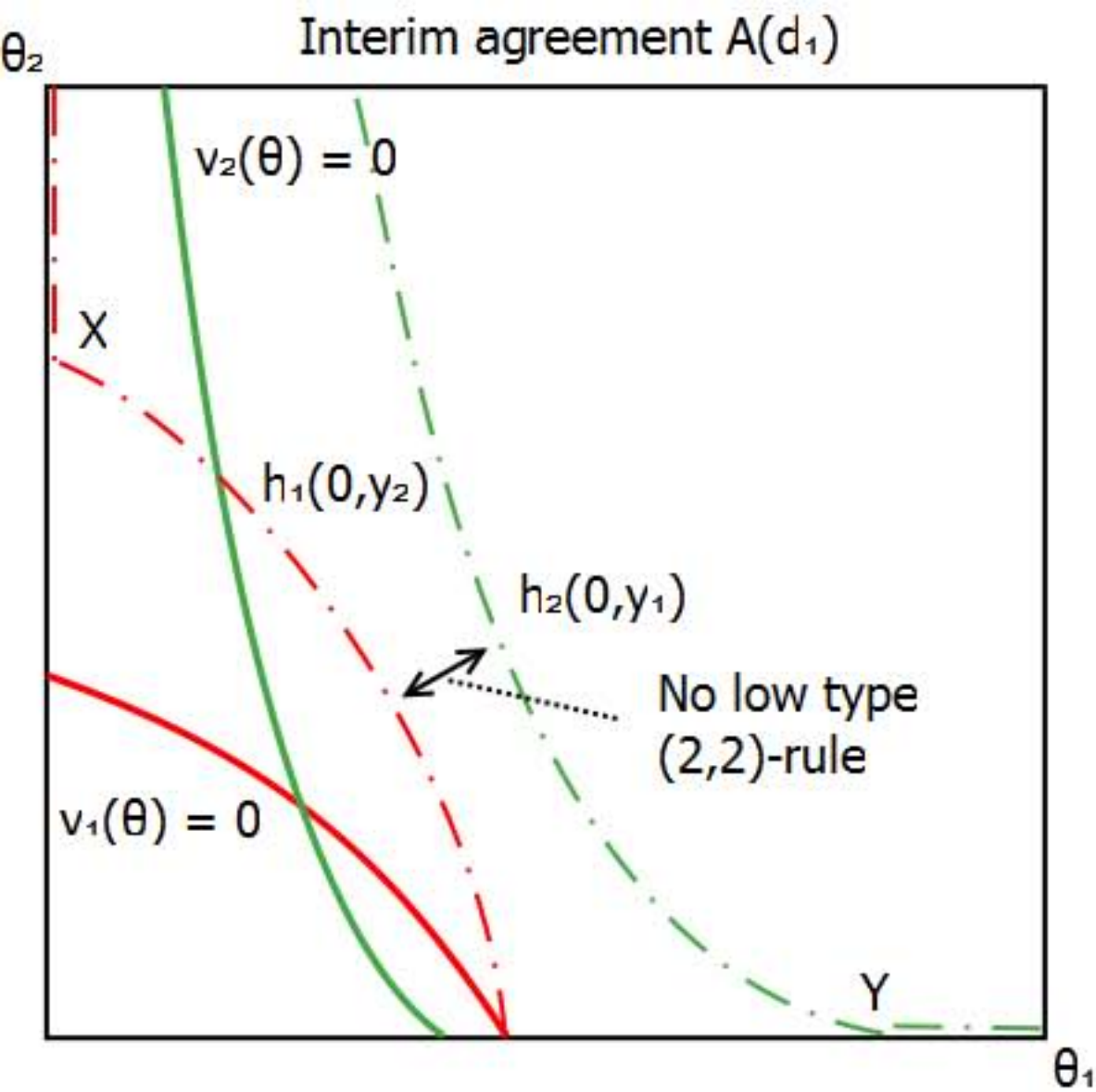
Low (2,2)-rule
is implementable
on $[a, a_2] \times [b, b_2]$

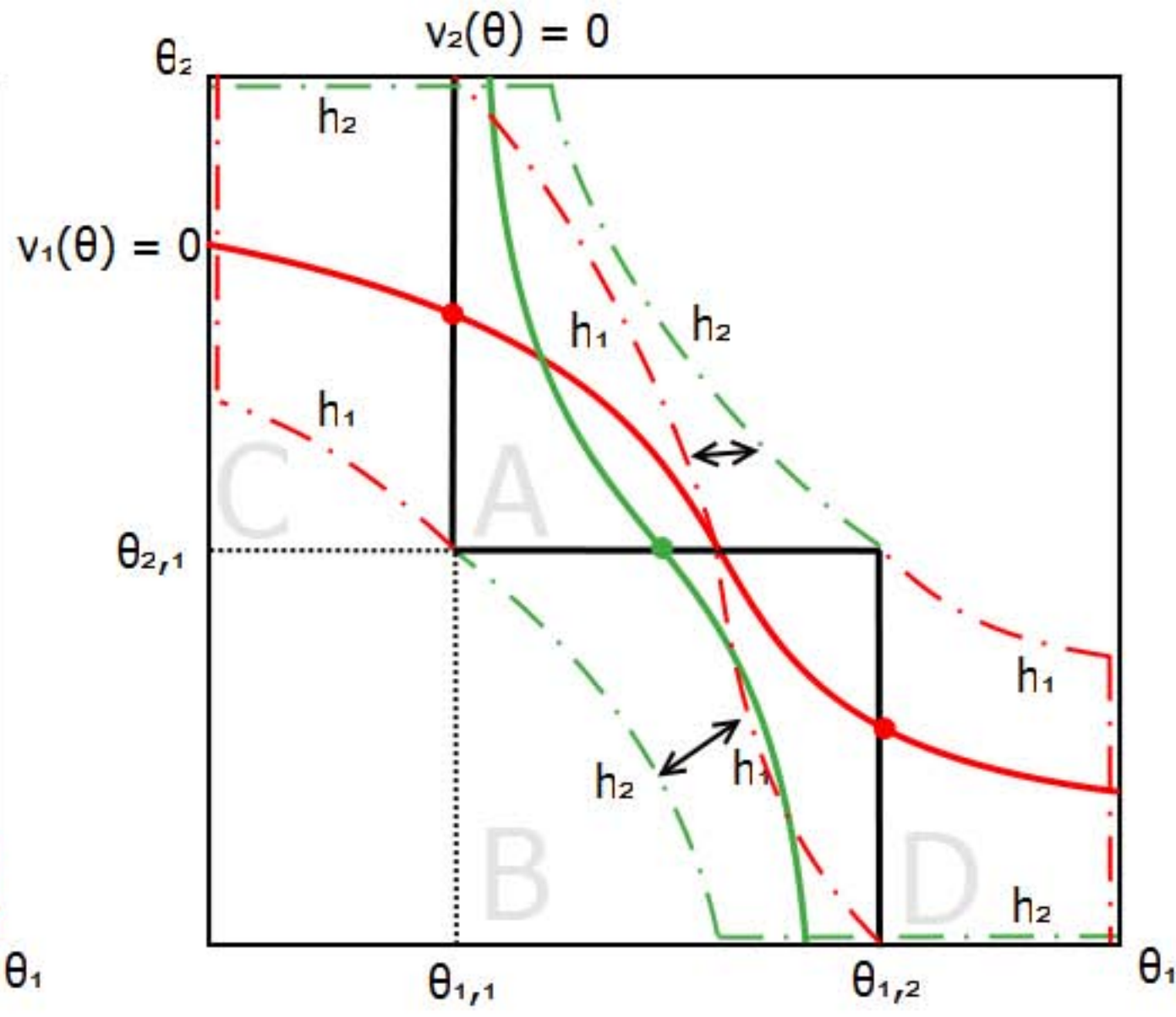
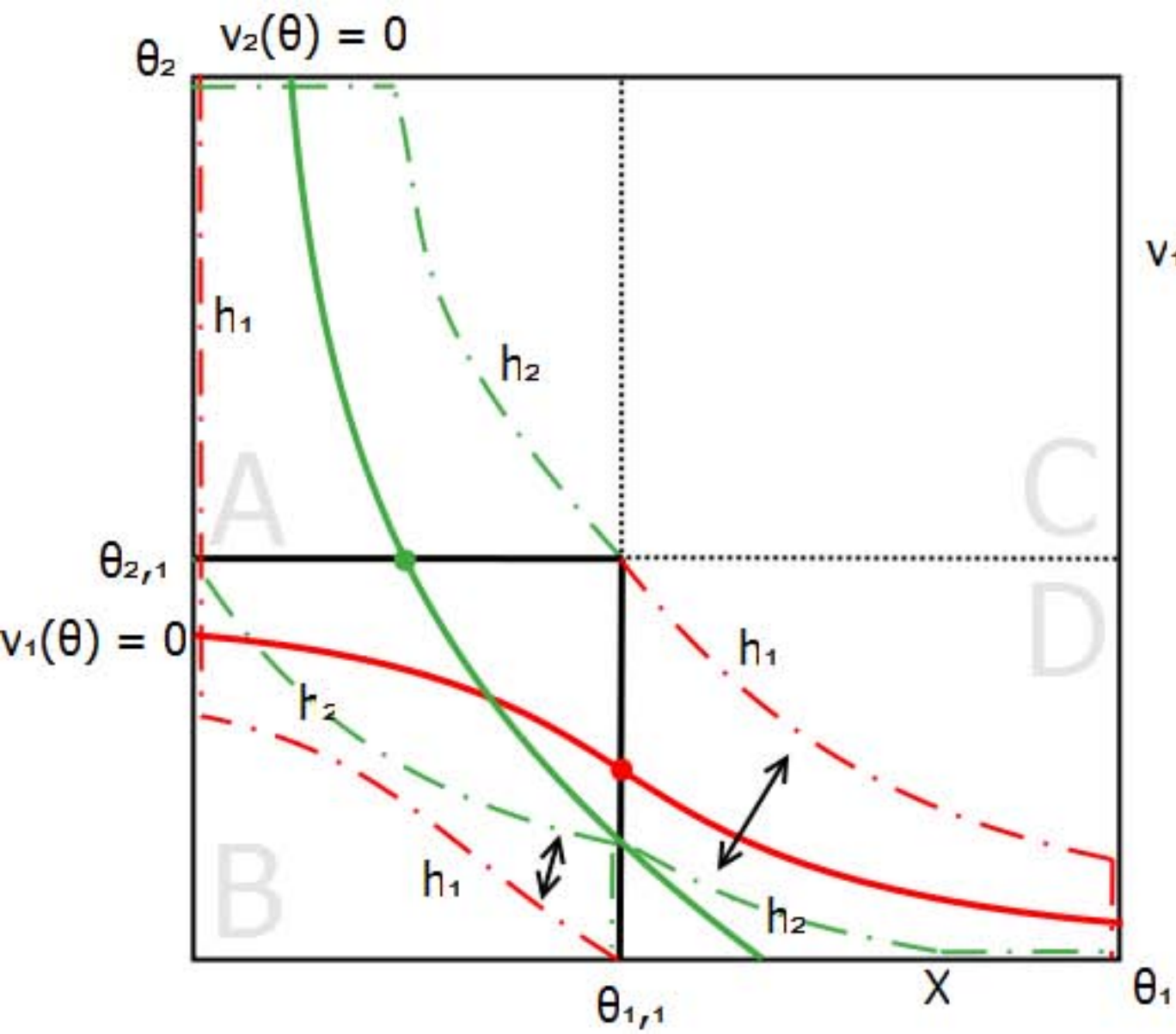
$$h_2(a, y_1)$$

$$v_1(\theta) = 0$$

 b_2 b a a_2 θ_1 









Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Kawakami, K

Title:

Posterior renegotiation-proofness in a two-person decision problem

Date:

2016-11

Citation:

Kawakami, K. (2016). Posterior renegotiation-proofness in a two-person decision problem. INTERNATIONAL JOURNAL OF GAME THEORY, 45 (4), pp.893-931. <https://doi.org/10.1007/s00182-015-0491-9>.

Persistent Link:

<http://hdl.handle.net/11343/282794>