The Pennsylvania State University

The Graduate School

College of the Liberal Arts

Essays in Game Theory and Political Economy

A Thesis in
Economics
by

Han Joon Jung

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2007

The thesis of Han Joon Jung was reviewed and approved* by the following:

James Jordan

Professor of Economics

Thesis Adviser

Chair of Committee

Edward Green

Professor of Economics

Sophie Bade

Assistant Professor of Economics

Frank Dardis

Assistant Professor of Advertising/Public Relations

James Jordan

Professor of Economics

Department Head

*Signatures are on file in the Graduate School

Abstract

We study models about game theory and political economy. The first chapter titled “Strategic Information Transmission through the Media” studies strategic communication through the media and provides several simple models that demonstrate how a sender or senders can manipulate information through the news media, such as newspaper or television stations, in order to influence the decision making process of receivers. The second chapter titled “Spatial Pillage Game” studies a coalitional game as a literature on allocation by force. Spatial pillage game extends one of the pillage games which models Hobbesian anarchy by introducing a space concept into a pillage game. The third chapter studies a prisoner’s dilemma game in the random-matching setting without information flow across matches and examines the Possibility of cooperation.

Contents

CHAPTER 1

Strategic Information Transmission through the Media

The first chapter model media manipulation in which a sender or senders manipulate information through the media to influence the decisions of receivers. We show that if there is only one sender and the receivers are in a situation of potential conflict, the sender successfully influences the receivers to play the sender's favorite outcome even when the sender and the receivers have contradictory preferences. This result extends to the cases in which the sender has imperfect information or in which the sender most values its credibility in reporting accurate information. In the case of multiple senders, however, if a sender receives a sufficient reward for reporting truthfully when others do not, all senders have incentives to report truthfully. Consequently, the receivers could play their favored outcome against the senders' preferences.

1.1. Introduction

With rapid development of technology, the mass media have become an influential force in our daily lives as a means of information transmission. Without a systematic framework to analyze the effect of the media, we are unaware of the impact of media and are subject to their hidden yet powerful influence. This paper intends to make an initial contribution to the understanding of the systematic operation and the impact of the media. More specifically, this paper provides several simple models that demonstrate how a sender or senders can manipulate information through the news media, such as newspaper or television stations, in order to influence the decision-making process of receivers.

The setting of the models is an arms race game. In the arms race game, there are two players, player 1 and player 2, each simultaneously choosing to build weapons or not to build weapons. In this game, player 2 has two possible types, either hawkish or dovish. If player 2 is hawkish, she wants to occupy a leading military position and therefore regards building weapons as a dominant action. If player 2 is dovish, she wants to maintain a harmonious relationship with player 1. Thus, dovish player 2 prefers not to build weapons as long as player 1 does not build weapons. As a dovish player who also wants to defend herself, however, dovish player 2 prefers to build weapons if player 1 builds weapons. In other words, dovish player 2 wants to match the actions of player 1. Player 2's type is her own private information. On the other hand, player 1 is always dovish, so she always wants to match the action of player 2[1]. In this arms race game, there

[1] In the classic arms race game, both players have two possible types. By designating one player's type as permanent, I simplify the classic arms race game while preserving the same results as the classic models.

are two possible equilibrium outcomes: the building-weapons outcome in which both players choose to build weapons and the not-building-weapons outcome in which both players choose not to build weapons. Here, the not-building-weapons outcome is the favored outcome for dovish players.

In the basic model, I introduce a sender into the arms race game. The sender has the information about player 2's type and would report the information to players 1 and 2 before they make decisions on weapon building. The following three assumptions about the sender lead to the unique outcome of this game. First, the sender reports the information about player 2's type through news media. By the nature of the news media, the report from the sender is commonly known to both players. Second, the sender has a preference for player 1 to build weapons. Finally, the sender has a conditional preference[2] for maintaining its credibility in reporting accurate information. Whether a sender can successfully influence the players or not is determined by the players' own strategies. If player 1 uses a strategy to ignore the sender's report and consequently the sender cannot influence

[2]This model is a simplified version of the repeated game in which a sender has two possible types: neutral or biased. A neutral sender always reports truthfully, while on the other hand, the sole concern of a biased sender is to make player 1 build weapons. In this repeated game, in response to the behavior that player 1 tries to distinguish a neutral sender from a biased sender, a biased sender pretends to be neutral in order to preserve its influence on player 1. In the basic model, to reflect this behavior of the biased sender, the concept of credibility is adopted and adapted so that the sender considers its credibility only conditionally. For more information about credibility in static games, see Kartik, Ottaviani, and Squintani (2006).

player 1 to build weapons, the sender would choose to preserve its credibility and thus report truthful information about player 2's type. However, if the sender can successfully affect player 1 to build weapons, then maintaining its credibility in reporting accurate information is no longer a concern to the sender and it could report untruthful information. With this sender in the arms race game, the basic model[3] shows that the only outcome in the Perfect Bayesian equilibrium, introduced by Fudenberg and Tirole (1991), is the building-weapons outcome.

To see why the players cannot achieve the not-building-weapons outcome in equilibrium, suppose that player 1 tries not to build weapons regardless of the report from the sender. Under this strategy of player 1, the sender cannot influence player 1 to build weapons, so the sender would report truthful information about player 2's type. Then when the sender reports that player 2 is hawkish, player 1 is certain that player 2 is hawkish and thus would build weapons. Accordingly,

[3]This basic model can be exemplified by the Hitler's regime in the World War II. According to Shirer (1990), Hitler (the sender) planned to start a war with the Poles (player 2) while German citizens (player 1) openly objected to the war. Since Hitler could not go into the war with an objection of his people, he manipulated the information about Polish policy to Germany considering that German citizens were shut off from the outside world. The evening before German attack, on September 1, 1939, German government broadcasted Fuehrer's Polish peace proposal at all German radio stations. But in fact, Hitler had never presented the proposal to the Poles. After its announcement, the German government claimed that their peace proposal had been rejected. On the same day, there were several faked "Polish attacks" by German S.S. ruffian at the Polish border, under Hitler's direction. The news about these attacks was also broadcasted on the German radio station right after the attacks. These series of events made the German citizens feel that they had no choice but to engage in the war with the Poles, which showed that Hitler successfully manipulated German citizens' opinions through the news media.

player 1 has an incentive to change her action from not building weapons to building weapons in order to match player 2's action. Therefore, in equilibrium, player 1 cannot completely ignore the report from the sender. Once player 1 responds to the sender's report, the sender can influence player 1 to build weapons by manipulating the information.

The result derived from the present model shows how powerful news media is as a means of information transmission. By reporting through the news media, the sender can make player 1 build weapons in accordance with the sender's preference, and therefore players 1 and 2 lose the not-building-weapons outcome, which is a favorite outcome of the dovish players. In addition, this result is strong because it is robust against two parameters. The result is robust against the probability that player 2 is hawkish. It is also robust against the payoffs to the dovish players when they achieve the not-building-weapons outcome. That is, no matter how small, but positive, the probability of player 2's being hawkish or no matter how great the payoffs to the dovish players in the not-building-weapons outcome, players 1 and 2 cannot achieve the not-building-weapons outcome.

Moreover, this result differs from the results in existing literature on information transmission. Crawford and Sobel (1982) and Kartik, Ottaviani, and Squintani (2006) showed that if a sender and receivers have contradictory preferences, then the sender cannot influence the receivers to play the sender's favorite outcome (see also Milgrom, 1981; Sobel, 1985; and Krishna and Morgan, 2001). The basic model, on the other hand, shows that if receivers are in a situation of potential conflict, which is an arms race game in the model, and a sender reports its information through news media, then the sender can influence the receivers to play the sender's favorite outcome even when there are contradictory preferences between the sender and the receivers.

The setting of the model, the arms race game, is developed from Schelling (1960) and Baliga and Sjöström (2004). Schelling argued that reciprocal fear from surprise attack makes defensive action desirable and can cause a multiplier effect in which both sides generate and escalate negative expectations of their opponents. This multiplier effect induces arms race even though the probability of each side being hawkish is just a small positive number. Baliga and Sjöström formally modeled Schelling's insight. The aforementioned arms race differs from Baliga and Sjöström in that uncertainty lies only on one side and so there is no multiplier effect. Even so, the basic model still shows that an arms race is always triggered by the sender's will because the sender uses the media as a means of information transmission.

In this paper, reports are relevant to payoffs, and this property distinguishes this model from the cheap-talk games developed by Farrell and Gibbons (1989a), Farrell and Rabin (1996), and Battaglini (2002). In cheap-talk games, talk is irrelevant to payoffs. So, one player does not need to believe what other players talk about (see also Farrell and Gibbons, 1989b; Stein, 1989; Farrell, 1993; Baliga and Morris, 2002; and Aumann and Hart, 2003). In the present model, on the other hand, the payoffs to the sender depends on its report as well as on other players' actions. This is done by assuming that the sender wants to preserve its credibility in reporting accurate information. Ironically, this intention of the sender for its credibility deprives the dovish players of their favorite outcome. In addition, the present model differs from the signaling games developed by Spence (1973), Cho and Kreps (1987), and van Damme (1989) in that the sender reports player 2's type. In signaling games, senders signal their own types or their own intention about their future actions (see also Bhattacharya, 1979; Milgrom and Roberts, 1986; Banks and Sobel, 1987; Manelli, 1997).

Besley and Prat (2004) and Baron (2005) also studied media manipulation in which the senders manipulate information through the media to influence receivers. Besley and Prat modeled a situation in which media outlets maintain a cozy relationship with the government. In exchange for compensation, such as a direct monetary payment or beneficial regulations, the media suppress embarrassing information about the government. Also contributing to this topic, Baron employed information competition between two interest groups, each advocating their positions through news media to influence public sentiment. In both papers, the extent to which the media reports information determines the degree of influence that the media exert on the public, which actualizes media manipulation (see also Dyck and Zingales, 2003; and Strömberg, 2004). That is, more reports bring stronger influence on the public, and so the influence of the media on the decision of the public is exogenously modeled. In the present study, however, the influence of the media on the decision of the players is endogenously created in equilibrium due to the potential conflict between the receivers.

Section 3 extends the basic model by introducing imperfect information so that the sender detects imperfect information about player 2's type. Here, if the signal about player 2's type indicates the true type of player 2 with a sufficiently high probability, then the result in the basic model extends to the imperfect information model.

Section 4 examines the opposite case of the basic model in terms of the sender's preferences. In this modified model, the sender's primary preference is to preserve its credibility, and its conditional preference is to influence the decision-making process of players. This is done by assuming that the sender is composed of a private media outlet whose preferences on players' choices represent media bias. This assumption leads to the fact that as long as its credibility remains intact, a profit-maximizing media outlet manipulates the information to increase the demand for its products. This modified model concludes that the media outlet influences the players to play its favorite outcome without undermining its credibility.

Section 5 extends the modified model to incorporate private media competition to study the effect of media competition on media bias. This media competition model shows that if the profits of private media outlets are affected by media bias more than by media competition, then the media outlets can successfully manipulate the information so that players choose the outcome that satisfies media outlets' concerns. However, when there are both media outlets who report truthfully and those who report untruthfully, if competition among media outlets is strong enough so that competition brings sufficient rewards to the media outlets who report truthfully, then eventually all media outlets can be forced to report truthfully and receivers can achieve the not-building-weapons outcome. Therefore, enough competition among the private media outlets can effectively curb their information manipulation through the media and consequently reduce their influences on the receivers. Also, simultaneous reporting by at least two media outlets results in a outcome similar to the media competition case.

Section 6 presents summaries and conclusions.

1.2. Basic Model

In the basic model, a Ministry of Propaganda is the sender. There are three players; player 1, player 2, and the Ministry of Propaganda (M). Players 1 and 2 have their own types. Player 2 can be either hawkish or dovish while player 1 is always dovish. The probability that player 2 is hawkish is $h \in (0, 1]$. In this model, uncertainty lies only in the type of player 2. M reports either *that player* 2 *is hawkish* (H) or *that player* 2 *is dovish* (D). Players 1 and 2 each choose either *to Build weapons* (B) or *Not to build weapons* (N).

This game proceeds as follows. At stage zero, Nature chooses player 2's type. Only M and player 2 detect player 2's type. At stage one, M reports H or D. What M has reported becomes common knowledge. At stage two, players 1 and 2 each simultaneously choose B or N. After all actions are taken, payoffs are realized.

The payoffs to M depend on its own reports as well as the players' actions. That is, given a report $r \in \{H, D\}$ of M and actions $a_1, a_2 \in \{B, N\}$ of the players, a real number $u_{ra_1a_2}$ denotes the payoff to M when M reports r and players 1 and 2 choose a_1 and a_2, respectively. For example, u_{HBN} denotes the payoff to M when M reports H and players 1 and 2 choose B and N, respectively. While in the traditional information transmission models studied by Crawford and Sobel

(1982), Austen-Smith (1990), and Seidmann and Winter (1997), the payoffs to a sender or senders depend on receivers' actions only, in the present model, however, the payoffs to M depend on both its own actions and the players' actions.

Regarding its preferences, the primary preference of M is to make player 1 build weapons; *i.e.* $u_{rBa_2} > u_{r'Na'_2}$ for any $r, r' \in \{H, D\}$ and $a_2, a'_2 \in \{B, N\}$ in which the left side term denotes the payoff to M when player 1 builds weapons and the right side term denotes the payoff to M when player 1 does not. M might have a particular preference on player 2's actions. In this model, however, such a preference does not affect results as long as the primary preference of M is to make player 1 build weapons. So the preference of M on player 2's action is omitted.

In addition, when M is unable to influence player 1 to build weapons, the conditional preference of M is to preserve its credibility. In this model, there are two possible cases in which M might lose its credibility. In one case, M might lose its *Credibility related to Truthfulness* (CT). A hawkish player has a dominant action B. So if M has reported H and player 2 plays N, then player 1 is certain that M has lied, and thus M would lose its CT. Hence if M cannot affect player 1 to build weapons and expects player 2 to play N, then M prefers to choose D; *i.e.* $u_{DNN} > u_{HNN}$. In the other case, M might lose its *Credibility related to Accurate Warning* (CAW). If M expects player 2 to play B and reports D, then M would fail to warn player 1 of the danger that player 2 would build weapons and thus lose its CAW. Hence if M cannot affect player 1 to build weapons and expects player 2 to play B, then M prefers to report H; *i.e.* $u_{HNB} > u_{DNB}$.

However, not every inequality stated above to summarize the preferences of M affects the outcomes in equilibrium. Of the aforementioned inequalities, only the four listed below influence the outcomes. In this paper, emphasis is placed on the outcomes in equilibrium. Therefore, for simplicity, only the following four inequalities are assumed to describe *the preferences of the Ministry of Propaganda, M*;

i) $u_{HBB} > u_{DNN}$ and $u_{DBB} > u_{HNN}$ for the primary preference;

ii) $u_{DNN} > u_{HNN}$ for CT; and *iii)* $u_{HNB} > u_{DNB}$ for CAW.

The payoffs to players 1 *and* 2 are given by the following matrixes. In these matrixes, player 1 chooses a row and player 2 a column,

When player 2 is hawkish

	N	B
N	$\omega, 3$	$0, 4$
B	$2, 0$	$1, 1$

When player 2 is dovish

	N	B
N	ω, ω	$0, 2$
B	$2, 0$	$1, 1$

Table 1.1. Payoff Matrixes of Players 1 and 2

such that $\omega > 2$ where the first entry in each cell is player 1's payoff for the corresponding actions and the second entry player 2's. Player 1 is always dovish. A dovish player prefers to match the action of the other and prefers the NN outcome to the BB. For a hawkish player, B is a dominant action. So the BB outcome is the only pure-strategy equilibrium in the left side matrix. Moreover, when both players 1 and 2 are dovish, they want to match the action of the other. Consequently, in the right side matrix, there are two pure-strategy equilibria, the NN outcome and the BB. This two-player setting is similar to Baliga and Sjöström (2004).

The arms race game without M has two pure-strategy Perfect Bayesian equilibrium outcomes. Players can achieve the BB outcomes regardless of player 2's type. Also, the NB outcome when player 2 is hawkish and the NN when player 2 is dovish is a possible outcome combination in equilibrium. If the probability h that player 2 is hawkish is small or if the payoff ω in the NN outcome is large, then both players prefer the latter outcome combination to the former because the expected payoffs in the latter are greater than the expected payoffs in the former. Then, how does introducing M into the arms race game change the results? Theorem 1 answers this question.

Theorem 1. *Pure-strategy Perfect Bayesian equilibria exist, and in the equilibrium outcomes players* 1 *and* 2 *choose* Bs*.*

Proof. Here, every pure-strategy of player 1 is examined. First, let player 1 play B always, *i.e.* (B, B). Note that a dovish player prefers to match the action of the other player. On the other hand, a hawkish player has the dominant action B. Hence only player 2's strategy under which she always plays B, *i.e.* (B, B, B, B), satisfies the best response to player 1's strategy in each continuation game. Finally, player 1's strategy (B, B) satisfies the best response to (B, B, B, B) in each continuation game. In this case, if $u_{HBB} > u_{DBB}$, then M prefers to report H always, *i.e.* (H, H). Accordingly, the strategy profile $\{(H, H), (B, B), (B, B, B, B)\}$ is an equilibrium. Similarly, the following strategy profiles are equilibria; if $u_{HBB} < u_{DBB}$, then $\{(D, D), (B, B), (B, B, B, B)\}$; and if $u_{HBB} = u_{DBB}$, then $\{(r_H, r_D), (B, B), (B, B, B, B)\}$ for $r_H, r_D \in \{H, D\}$, where (r_H, r_D) specifies that M reports r_H when it detects a hawkish type and r_D when it detects a dovish type. In all the cases, players always choose Bs in the outcomes. Therefore, there exists a pure-strategy equilibrium of which players choose Bs in the outcome.

Second, let player 1 play B only when M has reported H, *i.e.* (B, N). Then only player 2's strategy under which she plays N only when player 2 is dovish and M has reported D, *i.e.* (B, B, B, N), satisfies the best response in each continuation game.

Next, the best response of M to these strategies is to report H when M detects a dovish type because $u_{HBB} > u_{DNN}$, which denotes the primary preference of M. In this case, first, let M report D when it detects a hawkish type. Then M would report D when it detects a hawkish type and report H when it detects a dovish type. So player 1 knows that player 2 is hawkish when M has reported D. Hence player 1 has an incentive to change her action from N to B when M has reported D. Consequently, the strategy profiles that contain player 1's strategy (B, N) and M's strategy under which M reports D only when it detects a hawkish type, *i.e.* (D, H), cannot be an equilibrium. Second, let M report H when it detects a hawkish type, then the player 1's strategy (B, N) satisfies the best response in each continuation game. Hence players choose Bs in this outcome.[4] Therefore, if a strategy profile in which player 1 plays the strategy (B, N) is an equilibrium, then players choose Bs in the outcome of this equilibrium.

Third, let player 1 play N only when M has reported H, *i.e.* (N, B). Then only player 2's strategy under which she plays N only when player 2 is dovish and M has reported H, *i.e.* (B, B, N, B), satisfies the best response in each continuation game. Next, the best response of M to these strategies is to report D when it detects a dovish type because $u_{DBB} > u_{HNN}$, the primary preference of M. Similar to the

[4]If $u_{HBB} \geq u_{DNB}$ holds, then $\{(H, H), (B, N), (B, B, B, N)\}$ is a perfect bayesian equilibrium.

previous situation, if M takes the action H when it detects a hawkish type, then the strategy profiles in which player 1 plays (N, B) and M plays (H, D) cannot be an equilibrium. On the other hand, if M reports D when it detects a hawkish type, then players choose Bs in this outcome.[5] Therefore, if a strategy profile in which player 1 plays (N, B) is an equilibrium, then players choose Bs in this equilibrium outcome.

Finally, let player 1 play N always, *i.e.* (N, N). Then only player 2's strategy under which she plays N only when she is dovish, *i.e.* (B, B, N, N), satisfies the best response to player 1's strategy in each continuation game. Next, the best response of M is to report H when it detects a hawkish type because $u_{HNB} > u_{DNB}$, the conditional preference for CAW, and to report D when it detects a dovish type because $u_{DNN} > u_{HNN}$, the conditional preference for CT. Hence when M reported H, player 1 has an incentive to change her action from N to B because she is certain that player 2 is hawkish and thus player 2 will choose B. Therefore, the strategy profiles in which player 1 plays the strategy (N, N) cannot be an equilibrium. □

Theorem 1 means that only the BB outcomes are possible in pure-strategy perfect bayesian equilibrium and M successfully manipulates the information through

[5] If $u_{DBB} \geq u_{HNB}$ holds, then $\{(D, D), (N, B), (B, B, N, B)\}$ is a perfect bayesian equilibrium.

the news media. Therefore, introducing M into the arms race game lowers the players' payoffs. This result is strong in that it does not depend on h (> 0), the probability that player 2 is hawkish, and ω (> 2), the payoff to the dovish players in the NN outcome.

In contrast to the outcomes in pure-strategy equilibrium, the outcomes in mixed-strategy equilibrium can result in players choosing Ns with positive probabilities. However, a mixed-strategy equilibrium has negative features for the players. First, the expected payoffs in mixed-strategy equilibrium is relatively small compared with the expected payoffs in the combination of the NN outcome and the NB. If player 1 or player 2 is indifferent between playing B or N, then her expected payoff in that information set is $\frac{\omega}{\omega-1}$, which is less than 2, no matter which mixed strategy she plays. If ω is large enough, then $\frac{\omega}{\omega-1}$ is pretty small compared with ω. Second, the set of parameters that admit mixed-strategy equilibria has a small size under the following assumptions. Suppose that $u_{HBB} > u_{DBB}$ holds and the difference between u_{HBN} and u_{DBN}, or between u_{HNN} and u_{DNN} is large enough. Then the set of the parameters that admit mixed-strategy equilibria has measure zero.

1.3. Imperfect Information

The basic model is extended to the case in which M observes an imperfect signal about player 2's type. So in this extended model, instead of directly detecting player 2's type, M detects a signal that is exogenously given and is correlated with player 2's type. However, the signal can be empty. If the signal is empty, then it does not reveal any information about player 2's type. The probability that the signal indicates a hawkish type is p_{HH} when player 2 is hawkish and is p_{DH} when player 2 is dovish. The probability that the signal indicates a dovish type is p_{HD} when player 2 is hawkish and is p_{DD} when player 2 is dovish. Thus the probability that M detects an empty signal is $p_{HE} = 1 - p_{HH} - p_{HD}$ when player 2 is hawkish and is $p_{DE} = 1 - p_{DH} - p_{DD}$ when player 2 is dovish.

Here, M has three actions, H, D, and E. The action E denotes *the empty report.* We can regard E as no new information reported. In this extended model, M has one more option, E, than in the basic model. If M tries to report truthfully, it cannot report anything when it detects an empty signal. Therefore, this setting reflects the fact that M might report truthfully because of its conditional preference. The other settings are the same as in the basic model except for the interpretation of the conditional preference of M.

In this imperfect information model, M might report H by mistaking a dovish type for a hawkish type. Hence player 1 does not know whether or not M gave a wrong report on purpose. As a result, M might not lose its credibility related to truthfulness when it has given a wrong report. Even then, M would lose its *Credibility related to Accurate Forecast* (CAF) if it fails to correctly forecast the action of player 2. This is because a hawkish type has the dominant action B and by reporting H, M can forecast the action B. Accordingly, when M is unable to influence player 1 to build weapons, M prefers to report H only when it expects that player 2 would play B. In this model, therefore, the conditional preference of M is to preserve its CAF. The primary preference of M is still to make player 1 build weapons. Consequently, the following inequalities are assumed to describe *the preferences*[6] *of the Ministry of Propaganda,* M*, with imperfect information*;

i) $u_{HBB} > \max\{u_{DNB}, u_{DNN}, u_{ENB}, u_{ENN}\}$, $u_{DBB} > \max\{u_{HNB}, u_{HNN}, u_{ENB}, u_{ENN}\}$, and $u_{EBB} > \max\{u_{HNB}, u_{HNN}, u_{DNB}, u_{DNN}\}$ for the primary preference; and

ii) $u_{HNB} > \max\{u_{DNB}, u_{ENB}\}$ and $u_{DNN} > u_{HNN}$ for CAF.

Theorem 4 shows that if the signal about player 2's type is informative, then Theorem 1 extends to the imperfect information model. More specifically, if p_{HH}, the probability that the signal correctly indicates a hawkish type, and p_{DD}, the probability that the signal correctly indicates a dovish type, are high enough, then players 1 and 2 choose Bs in the equilibrium outcomes. Definitions 2 and 3 prescribe the levels of informativeness of the signal at which the signal could influence the decision-making process of M and player 1. Since player 2 knows her own type, the signal does not directly influence the decision of player 2. Hence

[6]For Theorem 4, it suffices to assume that $u_{HBB} > \max\{\frac{hs_{HD}u_{DNB}+(1-h)s_{DD}u_{DNN}}{hs_{HD}+(1-h)s_{DD}}, \frac{hs_{HD}u_{ENB}+(1-h)s_{DD}u_{ENN}}{hs_{HD}+(1-h)s_{DD}}, \frac{hs_{HE}u_{DNB}+(1-h)s_{DE}u_{DNN}}{hs_{HE}+(1-h)s_{DE}}, \frac{hs_{HE}u_{ENB}+(1-h)s_{DE}u_{ENN}}{hs_{HE}+(1-h)s_{DE}}\}$ instead of $u_{HBB} > \max\{u_{DNB}, u_{DNN}, u_{ENB}, u_{ENN}\}$; to assume that $u_{DBB} > \max\{\frac{hs_{HD}u_{HNB}+(1-h)s_{DD}u_{HNN}}{hs_{HD}+(1-h)s_{DD}}, \frac{hs_{HD}u_{ENB}+(1-h)s_{DD}u_{ENN}}{hs_{HD}+(1-h)s_{DD}}, \frac{hs_{HE}u_{HNB}+(1-h)s_{DE}u_{HNN}}{hs_{HE}+(1-h)s_{DE}}, \frac{hs_{HE}u_{ENB}+(1-h)s_{DE}u_{ENN}}{hs_{HE}+(1-h)s_{DE}}\}$ instead of $u_{DBB} > \max\{u_{HNB}, u_{HNN}, u_{ENB}, u_{ENN}\}$; and to assume that $u_{EBB} > \max\{\frac{hs_{HD}u_{HNB}+(1-h)s_{DD}u_{HNN}}{hs_{HD}+(1-h)s_{DD}}, \frac{hs_{HD}u_{DNB}+(1-h)s_{DD}u_{DNN}}{hs_{HD}+(1-h)s_{DD}}, \frac{hs_{HE}u_{HNB}+(1-h)s_{DE}u_{HNN}}{hs_{HE}+(1-h)s_{DE}}, \frac{hs_{HE}u_{DNB}+(1-h)s_{DE}u_{DNN}}{hs_{HE}+(1-h)s_{DE}}\}$ instead of $u_{EBB} > \max\{u_{HNB}, u_{HNN}, u_{DNB}, u_{DNN}\}$. These stronger assumptions are made for simplicity. However, they are innocuous in that they match the preferences of M, whose primary concern is to make player 1 build weapons.

the conditions in Definitions 2 and 3 are sufficient to support a result similar to Theorem 1.

Definition 2 exhibits the *levels of informativeness of the signal to M*. The informativeness of the signal is evaluated based on the payoffs to M.

Definition 2. *Suppose that $u_{HNB} > \max\{u_{DNB}, u_{ENB}\}$ and $u_{DNN} > u_{HNN}$, CAF. Then the signal is said to be* ***informative to M*** *if the probabilities of the signals satisfy the following two inequalities;*

$$(1.1)\quad \frac{hp_{HH}u_{HNB} + (1-h)p_{DH}u_{HNN}}{hp_{HH} + (1-h)p_{DH}} > \max\{\frac{hp_{HH}u_{DNB} + (1-h)p_{DH}u_{DNN}}{hp_{HH} + (1-h)p_{DH}}, \frac{hp_{HH}u_{ENB} + (1-h)p_{DH}u_{ENN}}{hp_{HH} + (1-h)p_{DH}}\} \text{ and}$$

$$(1.2)\quad \frac{hp_{HD}u_{HNB} + (1-h)p_{DD}u_{HNN}}{hp_{HD} + (1-h)p_{DD}} < \frac{hp_{HD}u_{DNB} + (1-h)p_{DD}u_{DNN}}{hp_{HD} + (1-h)p_{DD}}.$$

Suppose that player 1 plays the strategy that specifies that she always plays N, i.e. (N, N, N), and that player 2 plays the strategy that specifies that she plays B only when she is hawkish, *i.e.* (B, B, B, N, N, N). In this case, M cannot influence player 1 to build weapons. If M fails to correctly forecast the action of player 2, then M gains nothing but loses its CAF. So if the signals indicate the true type of player 2 with significantly high probabilities, then M would truthfully report what it detects. Given the payoffs to M, inequalities 1.1 and 1.2 provide precise levels of probabilities with which M prefers to report truthfully. More concretely, inequality 1.1 shows that M prefers to report H when it detects a hawkish type and inequality 1.2 shows that M prefers not to report H when it detects a dovish type.

Definition 3 formulates the *levels of informativeness of the signal to player* 1. Similar to the case of M, the informativeness of the signal is evaluated based on the payoffs to player 1.

Definition 3. *The signal is said to be* ***informative to player*** **1** *if the probabilities of the signals satisfy the following two inequalities;*

$$\frac{hp_{HH} + 2(1-h)p_{DH}}{hp_{HH} + (1-h)p_{DH}} > \frac{\omega(1-h)p_{DH}}{hp_{HH} + (1-h)p_{DH}} \quad \textit{and} \tag{1.3}$$

$$\frac{h(p_{HH} + p_{HE}) + 2(1-h)(p_{DH} + p_{DE})}{h(p_{HH} + p_{HE}) + (1-h)(p_{DH} + p_{DE})} > \frac{\omega(1-h)(p_{DH} + p_{DE})}{h(p_{HH} + p_{HE}) + (1-h)(p_{DH} + p_{DE})}. \tag{1.4}$$

First, let M report H only when it detects the signal of a hawkish type. Then inequality 1.3 provides precise levels of probabilities with which player 1 prefers to play B when M has reported H. Second, let M report H only when it detects either the signal of a hawkish type or an empty signal. Then inequality 1.4 provides precise levels of probabilities with which player 1 prefers to play B when M has reported H.

Theorem 4. *Pure-strategy Perfect Bayesian equilibria exist, and if the signal is informative to both M and player 1, then in the equilibrium outcomes players 1 and 2 choose Bs.*

Proof. It is easily seen that there exists an equilibrium that contains player 1's strategy (B, B, B) and player 2's strategy (B, B, B, B, B, B), which specify that they always play Bs. In the outcomes of this equilibrium, players play Bs. Therefore, it suffices to show that if a strategy profile is an equilibrium, then players play Bs in the outcomes of the strategy profile.

Let player 1 play N only when M has reported E, *i.e.* (B, B, N). Then only player 2's strategy under which she plays N only when she is dovish and M has reported E, *i.e.* (B, B, B, B, B, N), satisfies the best response in each continuation game. Note that if M reports E, then its expected payoff is a weighted average between u_{ENB} and u_{ENN}. Thus the best response of M to the players' strategies is not to report E no matter what M detects because $u_{HBB} > \max\{u_{ENB}, u_{ENN}\}$, the primary preference of M. Finally, player 1's strategy (B, B, N) satisfies the best response in each continuation game. Therefore, players choose Bs in this outcome. Similarly, players choose Bs in the equilibrium outcomes in which player 1 plays the strategies (B, N, B), (B, N, N), (N, B, B), (N, B, N), or (N, N, B).

Finally, let player 1 play (N, N, N). Then only player 2's strategy under which she plays N only when she is dovish, *i.e.* (B, B, B, N, N, N), satisfies the best response in each continuation game. Under these strategies, M prefers to report H when it detects a hawkish type and prefers not to report H when it detects a dovish type because the signal is informative to M. Then player 1 has an incentive to change her action from N to B when M has reported H, because the signal is also informative to player 1. Therefore, the strategy profiles in which player 1 plays the strategy (N, N, N) cannot be an equilibrium. □

Corollary 5. *There exists* $\varepsilon > 0$ *such that if* $\min\{p_{HH}, p_{DD}\} > 1 - \varepsilon$*, then the outcomes in pure-strategy Perfect Bayesian equilibrium are that players* 1 *and* 2 *always choose* Bs *while* M *reports* H*,* D*, or* E*.*

Proof. Since $p_{DH} \leq 1 - p_{HH}$ and $p_{HD} \leq 1 - p_{DD}$, this result directly follows from Theorem 4 and inequalities 1.1, 1.2, 1.3, and 1.4. □

We have examined and extended the basic model in which sender's primary preference is to influence players' choices, such as by making player 1 build weapons, and its conditional preference is to preserve its credibility. In the next section, the opposite case will be considered in which sender's primary preference is to preserve its credibility and its conditional preference is to influence players' choices. This is done by assuming that a private media outlet who values its credibility most has its own preferences on players' choices and so the private media outlet tries to manipulate information.

1.4. Media Bias

In this modified model, a private media outlet is the sender itself, and player 1 is the main audience to this media outlet. So, there are three players; player 1, player 2 and the private media outlet (M). The other settings are the same as in the basic model except for the payoffs to M.

The payoffs to M depend on its reports and the action of player 2. That is, given a report $r \in \{H, D\}$ of M and an action $a_2 \in \{B, N\}$ of player 2, a real number u_{ra_2} denotes the payoff to M when M reports r and player 2 chooses a_2. This setting reflects the fact that M is mainly an information provider and the information from M would be evaluated in conjunction with the action of player 2. Thus the payoff to M is not affected by the action of player 1. This payoff setting for the private media outlet, M, could be considered the counterpart of the payoff setting for the Ministry of Propaganda in the basic model in terms of the action of player 1, since the payoff to the Ministry of Propaganda is most affected by the action of player 1.

Furthermore, in contrast to the Ministry of Propaganda, the primary preference of the private media outlet, M, is to preserve its credibility. Just like in the basic model, there are two possible cases in which M might lose its credibility; it may lose its *Credibility related to Truthfulness* (CT) or its *Credibility related to Accurate Warning* (CAW). Gentzkow and Shapiro (2006) studied the behavior of the media outlets who prefer to provide accurate information, and they presented empirical evidence of such behavior. They also showed that the importance of credibility in media markets has been emphasized by the management of media outlets. For example, Heyward (2004), president of CBS News, remarked, "Nothing is more important to CBS than our credibility" (see also Kirkpatrick and Fabrikant, 2003; and Rather, 2004).

When M can preserve its credibility, M is assumed to prefer the case in which it correctly warns player 1 of the danger of player 2 than the case in which it correctly predicts peace, *i.e.* $u_{HB} > u_{DN}$. This preference represents *media bias*. This assumption about the media bias reflects the following intuition: player 2's developing new weapons *i*) makes player 1 feel more insecure; so *ii*) induces player 1, who is the main audience to M, to pay more attention to M; and thus *iii*) eventually increases the profit of M by expanding the demand for the products of M. This kind of assumption about media bias is not new in economic literature.

Mullainathan and Shleifer (2005) modeled media bias based on a similar assumption in which profit-maximizing media outlets slant news stories to increase the demand for their products.

Formally, the following inequalities are assumed to describe *the preferences of the private media outlet, M*;

$i)$ $\min\{u_{DB}, u_{DN}\} > u_{HN}$ for CT; $ii)$ $u_{HB} > u_{DB}$ for CAW; and

$iii)$ $u_{HB} > u_{DN}$ for the media bias.

with these preferences of M, we can derive a result similar to Theorem 1.

Theorem 6. *The unique outcome in pure-strategy Perfect Bayesian equilibrium is that M reports H and players 1 and 2 choose Bs. If u_{HN} is small enough, there is no mixed strategy equilibrium except pure-strategy equilibrium.*

Proof. The proof of the first assertion is omitted because it is similar to the proof of Theorem 1.

To prove the second assertion, let β be the probability with which M reports H when it detects a dovish type. Also, let a_2 and b_2 be the probabilities with which dovish player 2 chooses B when M has reported H and chooses B when M has reported D, respectively. Finally, Let

$$a_2' \equiv \frac{b_2 u_{DB} + (1-b_2)u_{DN} - u_{HN}}{u_{HB} - u_{HN}} \text{ and}$$
$$a_2'' \equiv \frac{\beta(1-h)(\omega-2) - h}{\beta(1-h)(\omega-1)}.$$

If a_2 is equal to a_2', M is indifferent between playing H or D when it detects a dovish type. Also, if a_2 is equal to a_2'', player 1 is indifferent between playing B or N when M has reported H. However, given payoff parameters and h, if u_{HN} is small enough, then a_2' is greater than a_2'' for any $b_2, \beta \in [0,1]$.

Let a_1 be the probability with which player 1 chooses B when M has reported H. First, if $a_2 < a_2'$, then M prefers to play D when it detects a dovish type. Note that M prefers to play H when it detects a hawkish type because $u_{HB} > u_{DB}$, the conditional preference for CAW. Then player 1 knows that player 2 is hawkish when M has reported H. Hence, player 1 prefers to play $a_1 = 1$. Second, if $a_2'' < a_2$, then player 1 prefers to play $a_1 = 1$. If u_{HN} is small enough, then $a_2 < a_2'$ or $a_2'' < a_2$ because $a_2'' < a_2'$. Hence player 1 would play only $a_1 = 1$, and thus player 2 of a dovish type would also play only $a_2 = 1$. Then M would report only H because $u_{HB} > u_{DB}$, the conditional preference for CAW, and $u_{HB} > u_{DN}$, the media bias. Therefore, there is no mixed-strategy equilibrium except pure-strategy equilibrium. $\square$

Theorem 6 shows that even when the primary preference of the sender is its credibility, the receivers do not benefit more than they do in the basic model. This is because the availability of information from M removes the possibility for players to choose Ns. Consequently, M's warning to player 1 induces both players to choose Bs.

From M's point of view, the worst payoff is when it is proved to have lied. Only the payoff parameter u_{HN} denotes this worst payoff. So, lower u_{HN} means that M values its credibility related to truthfulness, CT, more. Theorem 6 implies that more value placed on the sender's CT could result in worse outcomes for the receivers by removing a mixed-strategy equilibrium. Also, the payoff u_{HN} could be considered to be the punishment that player 1 inflicts on M when M has verifiably lied. This is because player 1 is M's main audience and so the payoffs to M depend largely on player 1's interest in M's information. Then Theorem 6 implies that player 1 cannot improve her expected payoff by punishing M when M has verifiably lied and sometimes player 1 could even lower her expected payoff by punishing M.

DeFleur and Ball-Rokeach (1989) and Morris and Shin (2002) also analyzed the effects of information from the media (see also Gamson, Croteau, Hoynes, and Sasson, 1992; and Bernhardt, Krasa, and Polborn, 2006). They assumed that the degree of influence that the media exert on the public is determined by the extent to which the media reports its information, and thus the influence of the media on the public was exogenously modeled. In the present model, in contrast to their works, the structure of the model endogenously creates the influence of the media on the players' decision-making process.

We have seen two kinds of information transmission models: the basic model and the media bias model. In both models, there is only one sender, and the sender achieves its favorite outcome by manipulating information through the news media. Then, what would occur if there are more than one sender? If there are more than one sender, the senders could compete with one another. Does competition matter in these information transmission models? Does reporting simultaneously or reporting sequentially make a difference? The next section answers these questions.

1.5. Media Competition

In this model, I extend the previous model by introducing another private media outlet. Therefore, there are four players; player 1, player 2, media outlet 1 (M_1), and media outlet 2 (M_2). Player 1 is still the main audience to both M_1 and M_2. The media outlets report sequentially or simultaneously. When they report sequentially, M_1 reports first. The other settings are the same as the media bias model except the payoffs to the media outlets.

The payoffs to each outlet depend on the reports from the other outlet as well as its own reports and the action of player 2. That is, given reports $r_1, r_2 \in \{H, D\}$ of the media outlets and an action $a_2 \in \{B, N\}$ of player 2, a real number $u^i_{r_1 r_2 a_2}$ denotes the payoff to M_i for $i \in \{1, 2\}$ when M_1 reports r_1, M_2 reports r_2, and player 2 chooses a_2. This setting allows us to examine the effect of media competition.

Regarding their preferences, just like in the media bias model, the primary preference of the media outlets is to preserve their credibility: CT and CAW. In this model, the conditional preference of the media outlets depends on both *media bias* and *media competition*. The media outlets have the same bias as in the media bias model. So the media outlets prefer the case in which they correctly warn player 1 of the danger of player 2 than the case in which they correctly predict peace. The following inequalities represent all possible cases of media bias; $\min\{u^1_{HHB}, u^1_{HDB}\} > \max\{u^1_{DHN}, u^1_{DDN}\}$ and $\min\{u^2_{HHB}, u^2_{DHB}\} > \max\{u^2_{HDN}, u^2_{DDN}\}$. Media competition imposes another condition for the payoffs to the media outlets by tempting them to occupy a leading position in competition. That is, in competition, media outlets have more incentive to predict correctly while others predict incorrectly than while others also predict correctly. The following inequalities describe all possible cases of media competition; $\min\{u^1_{DHN}, u^1_{HDB}\} > \max\{u^1_{HHB}, u^1_{DDN}\}$ and $\min\{u^2_{HDN}, u^2_{DHB}\} > \max\{u^2_{HHB}, u^2_{DDN}\}$.

However, the following two inequalities for media competition $u^1_{DHN} > u^1_{HHB}$ and $u^2_{HDN} > u^2_{HHB}$ directly contradict the inequalities for media bias $u^1_{DHN} < u^1_{HHB}$ and $u^2_{HDN} < u^2_{HHB}$. Therefore, there are two possible cases: one case in which bias dominates competition and the remaining case in which bias does not dominate competition. *When bias dominates competition*, the following inequalities are assumed to describe *the preferences of the private media outlets*;

i) $\min\{u^2_{HDB}, u^2_{HDN}\} > u^2_{HHN}$ for CT;

ii) $u^1_{HHB} > \max\{u^1_{DHB}, u^1_{DDB}\}$ and $u^2_{HHB} > u^2_{HDB}$ for CAW; and

iii) $u^1_{HHB} > \max\{u^1_{DDN}, u^1_{DHN}\}$ and $u^2_{HHB} > u^2_{HDN}$ for the media bias.

When bias does not dominate competition, the following inequalities are assumed to describe *the preferences*[7] *of the private media outlets*;

i) $u^1_{DDN} > u^1_{HDN}$ and $u^2_{DDN} > u^2_{DHN}$ for CT;

ii) $u^1_{HHB} > \max\{u^1_{DHB}, u^1_{DDB}\}$ and $u^2_{HHB} > u^2_{HDB}$ for CAW; and

iii) $u^2_{HDN} \geq u^2_{HHB}$ for the weakly dominant media competition.

---

[7] For Theorem 7, it suffices to assume that $u^2_{DDN} \geq u^2_{DHN}$ and $u^1_{DDN} \geq u^1_{HDN}$ instead of $u^2_{DDN} > u^2_{DHN}$ and $u^1_{DDN} > u^1_{HDN}$; and to assume that $u^1_{HHB} \geq \max\{u^1_{DHB}, u^1_{DDB}\}$ and $u^2_{HHB} \geq u^2_{HDB}$ instead of $u^1_{HHB} > \max\{u^1_{DHB}, u^1_{DDB}\}$ and $u^2_{HHB} > u^2_{HDB}$. These stronger assumptions are made in accordance with the preferences about credibility.

Note that when bias dominates competition, $u^2_{HHB} > u^2_{HDN}$ is assumed. On the other hand, when bias does not dominate competition, $u^2_{HHB} \leq u^2_{HDN}$ is assumed.

Theorem 7. *Suppose that the media outlets report sequentially. If bias dominates competition, then the unique outcome in pure-strategy Perfect Bayesian equilibrium is that both outlets report Hs and players choose Bs. In addition, if u^2_{HHN} is small enough, there is no mixed-strategy equilibrium except pure-strategy equilibrium. However, if bias does not dominate competition, then there exists a pure-strategy Perfect Bayesian equilibrium in which the media outlets report Hs and players choose Bs when player 2 is hawkish and the media outlets report Ds and players choose Ns when player 2 is dovish. Suppose that the media outlets report simultaneously and bias dominates competition. Moreover, add additional assumption that $u^1_{DDN} > u^1_{HDN}$ and $u^2_{DDN} > u^2_{DHN}$, for CT. Then there exists a pure-strategy Perfect Bayesian equilibrium in which the media outlets report Hs and players choose Bs when player 2 is hawkish and the media outlets report Ds and players choose Ns when player 2 is dovish.*

Proof. First, suppose that the media outlets report sequentially. It is easily seen that there exists an equilibrium outcome in which players 1 and 2 play Bs. Therefore, it suffices to show that if a strategy profile is an equilibrium, then players play Bs in the outcomes of the strategy profile.

Let player 1 play B when both outlets have reported Hs. Then the best response of player 2 to this action is to play B when both outlets have reported Hs. Next, the best response of M_2 to these actions is to play H when M_1 has reported H because $u^2_{HHB} > u^2_{HDB}$, CAW, and $u^2_{HHB} > u^2_{HDN}$, the media bias. Also, the best response of M_1 to these actions is to play H because $u^1_{HHB} > \max\{u^1_{DHB}, u^1_{DDB}\}$, CAW, and $u^1_{HHB} > \max\{u^1_{DDN}, u^1_{DHN}\}$, the media bias. Finally, the best response of player 1 is to play B when both outlets have reported Hs. Therefore, in this outcome, both outlets report Hs and players choose Bs.

Let player 1 play N when both outlets have reported Hs. Then the best response of dovish player 2 is to play N when both outlets have reported Hs. Next, the best response of M_2 is to play D when it detects a dovish type and M_1 has reported H because $\min\{u^2_{HDB}, u^2_{HDN}\} > u^2_{HHN}$, CT. Note that if the media outlets detect a hawkish type, then they know that player 2 would play B, and thus they prefer to report Hs because $u^2_{HHB} > u^2_{HDB}$ and $u^1_{HHB} > \max\{u^1_{DHB}, u^1_{DDB}\}$, CAW. Hence, player 1 can be certain that player 2 is hawkish when both outlets have reported Hs. Consequently, player 1 has an incentive to change her action from N to B when both outlets have reported Hs. Therefore, a strategy profile in which player 1 plays N when both outlets have reported Hs cannot be an equilibrium. This completes the proof of the first assertion.

The proof of the second assertion is omitted because it is similar to the proof in Theorem 6.

To prove the third and the fourth assertions, let player 1 play B only when both outlets have reported Hs, *i.e.* (B, N, N, N). Then only player 2's strategy under which she plays B either when she is hawkish or when both outlets have reported Hs, *i.e.* (B, B, B, B, B, N, N, N), satisfies the best response in each continuation game.

Suppose that bias does not dominate competition. Then when M_2 detects a dovish type, one of the best responses of M_2 is to report D no matter what M_1 has reported because $u^2_{DDN} > u^2_{DHN}$, CT, and $u^2_{HDN} \geq u^2_{HHB}$, the weakly dominant media competition. Next, when M_1 detects a dovish type, the best response of M_1 is to report D because $u^1_{DDN} > u^1_{HDN}$, CT. Note that if M_1 detects a hawkish type, then it would report H because $u^1_{HHB} > \max\{u^1_{DHB}, u^1_{DDB}\}$, CAW. Also, if M_1 has reported H and M_2 detects a hawkish type, then M_2 would report H because $u^2_{HHB} > u^2_{HDB}$, CAW. Under these strategies of player 2, M_1, and M_2, player 1's strategy (B, N, N, N) satisfies the best response in each continuation game. Therefore, there exists an action $r^2_{HD} \in \{H, D\}$ of M_2 such that the strategy profile $\{(H, D), (H, r^2_{HD}, D, D), (B, N, N, N), (B, B, B, B, B, N, N, N)\}$ is an equilibrium, where r^2_D specifies the action of M_2 when it detects a hawkish type and M_1 has reported D. In this strategy profile, the media outlets report Hs and players choose Bs when player 2 is hawkish and the media outlets report Ds and players choose Ns when player 2 is dovish.

Finally, suppose that the media outlets report simultaneously and bias dominates competition. Also, assume that $u^1_{DDN} > u^1_{HDN}$ and $u^2_{DDN} > u^2_{DHN}$ for CT. Then when the outlets detect a hawkish type, each outlet prefers to report H if the other outlet would report H because $u^1_{HHB} > u^1_{DHB}$ and $u^2_{HHB} > u^2_{HDB}$, CAW. Also, when the outlets detect a dovish type, each prefers to report D if the other would report D because $u^1_{DDN} > u^1_{HDN}$ and $u^2_{DDN} > u^2_{DHN}$, CT. Finally, player 1's strategy (B, N, N, N) satisfies the best response in each continuation game. Therefore, the strategy profile $\{(H, D), (H, D), (B, N, N, N), (B, B, B, B, B, N, N, N)\}$ is a pure-strategy equilibrium in which the media outlets report Hs and players choose Bs when player 2 is hawkish and the media outlets report Ds and players choose Ns when player 2 is dovish. □

Therefore, as long as bias dominates competition and the media outlets report sequentially, players 1 and 2 cannot improve their expected payoffs. Just like in the media bias model, the punishment for an untruthful report does not improve players' expected payoffs, and sometimes even lowers their payoffs. However, media competition and simultaneous reporting each can make the media outlets report truthfully, and consequently players can improve their expected payoffs. In fact, the outcome combination in which players 1 and 2 play Bs when player 2 is hawkish and they play Ns when player 2 is dovish gives player 1 the best payoff out of combinations in which hawkish player 2 plays B, and player 1 can achieve this combination only when the media outlets reveal the information about player 2's type. Also, when there are more than one sender in the basic model, if the senders report simultaneously or if competition among the senders is strong enough, then players can achieve this outcome combination. Therefore, **enough competition and simultaneous reporting can each solve media manipulation problems** in which the senders manipulate the information through the media to influence the receivers.

The media competition model can be extended by introducing more media outlets and imperfect signals about player 2's type. In this case, if bias dominates competition, then we can derive a result similar to Theorems 4 and 7. That is, if the signal is informative to player 1 and the media outlets, then the unique outcome in pure-strategy equilibrium is that all outlets report Hs and players choose Bs. However, if competition dominates bias, then the result depends on media outlets' *herd behavior* in reporting, which was studied by Scharfstein and Stein (1990) and Banerjee (1992), and the result might differ from Theorems 4 and 7. Here, herd behavior in reporting means that media outlets just follow other outlets' actions regardless of their own information about player 2's type.

To see how the result can be affected by the herd behavior, suppose that player 1 plays N if at least one media outlet reports D. If competition dominates bias and there is no herd behavior, then this strategy of player 1 influences the media outlets to report truthfully. However, in the imperfect information case, when media outlets truthfully report and it turns out that only one outlet has reported D and the other all outlets have reported Hs, player 1 has an incentive to change her action from N to B. This is because the probability that player 2 is hawkish is much higher than the probability that player 2 is dovish. Since player 1 would not play her strategy suggested above, the media outlets lose an incentive to report truthfully. In this case, if there is herd behavior in reporting, then even when all media outlets except one outlet have reported Hs, the probability that player 2 is hawkish might not be high. This is because most of the outlets just followed previous reports and accordingly just a few media outlets reveal their information about player 2's type. Consequently, if the payoff ω of player 1 in the NN outcome is large enough, then player 1 can rationally play N even when all media outlets except one outlet have reported Hs. Therefore, player 1 can achieve her favorite outcome, NN. In this sense, media outlets' herd behavior in reporting is considered to improve player 1's expected payoff.

In this model, I assume only rational players and focus on the effect of media bias. On the other hand, a large literature about media bias assumes irrational players to a certain extent and focuses on the occurrence and persistence of media bias. That is, they explained how media bias happens and continues based on the irrational player assumption. Mullainathan and Shleifer (2005), Baron (2006), and Gentzkow and Shapiro (2006) studied media bias and examined whether media competition can reduce media bias. Baron explained that the occurrence and persistence of media bias can arise from the supply side by journalists who are willing to accept lower wages for the sake of their discretion. Mullainathan, Shleifer, Gentzkow, and Shapiro showed that media bias can, however, come from the demand side. They argued that profit-maximizing media outlets could cater to the preferences or prior beliefs of their audience, who can be considered as irrational players, to increase the demand for their products; this behavior represents media bias. Regarding media competition, Mullainathan, Shleifer, and Baron concluded that competition by itself may not be powerful enough to reduce media bias. On the other hand, Gentzkow and Shapiro found that competition among independently owned media outlets can lead to lower bias. In the present model, similar to Gentzkow and Shapiro, sufficiently strong competition can lead the media outlets to report truthfully.

1.6. Conclusions

Suppose that receivers are in a situation of potential conflict and a sender reports its information through the news media. Then even when the sender and the receivers have contradictory preferences, the sender can make receivers play the sender's favorite outcome. This result extends to the imperfect information case. That is, when the information includes noise, if the information is informative enough, then the sender can achieve its favorite outcome.

This result does not change even when the sender values more its credibility in reporting accurate information. This is because the availability of the information from the sender removes the possibility for the players to play Ns. However, if there are more than one sender, then the result can change. If competition among the senders is strong enough or if at least two senders report simultaneously, then the senders can be forced to report truthfully, and as a result, the receivers' expected payoffs can improve. Therefore, enough competition and simultaneous reporting can be a solution to the media manipulation problem.

CHAPTER 2

Spatial Pillage Game

The second chapter studies a coalitional game. A pillage game is a coalitional game that is meant to be a model of Hobbesian anarchy. The spatial pillage game introduces a spatial feature into the pillage game by assuming that players are located in regions. Players can travel from one region to another in one move and can form a coalition and combine their power only with players in the same region. A coalition has power only within its region. Under this spatial restriction, some members of a coalition can pillage less powerful coalitions without any cost. The feasibility of pillages between coalitions determines the dominance relation. Core, stable set, and farsighted core are adopted as alternative solution concepts. We study three models about game theory and political economy.

2.1. Introduction

Hobbesian anarchy is a state of society before a government ensuring property rights is organized. Without such an organization, no individuals are safe to secure their wealth. Individuals could be tempted to pillage others whenever possible and beneficial. Although a coalition could be formed to secure their wealth, some members of the coalition may still be tempted to betray the others and to take their wealth. Consequently, in Hobbesian anarchy, the possibility of the stable distribution of wealth is questionable.

A substantial amount of literature on *allocation by force* has been devoted to this possibility. Skaperdas (1992) showed that a cooperative outcome is possible in equilibrium if the probability of winning in conflict is sufficiently robust against each individual's action. Hirshleifer (1995) found the conditions to make Hobbesian anarchy stable. Also, Hirshleifer (1991), Konrad and Skaperdas (1998), and Muthoo (1991) studied the situations in which property right is partially secured. These studies analyzed noncooperative models in which the formation of coalitions is limited or not allowed.

Different from the previous models, Piccione and Rubinstein (2006) and Jordan (2005) developed models of Hobbesian anarchy that allow the formation of coalitions. Piccione and Rubinstein introduced *the jungle* in which coercion governs economic transactions and they compared the equilibrium allocation of the jungle with the equilibrium allocation of an exchange economy. Jordan introduced *pillage games* and examined stable sets of allocations in which the power of pillaging balances endogenously.

The *spatial pillage game* is an extended version of a pillage game. In most literature on "allocation by force" including the papers reviewed above, there is no restriction on using power. Thus any individual or coalitions can pillage another individual or other coalitions if one is more powerful than others. However, the acts of pillaging and defending are inevitably under spatial restriction. Members of a coalition, if they move together, cannot simultaneously pillage two less powerful coalitions that are far apart from each other. Likewise, two coalitions cannot combine their power to defend themselves together against another powerful coalition unless they are close enough to each other. The spatial pillage game introduces a *space concept*, which conditions power usage based on location, into a Hobbesian anarchy model that allows the formation of coalitions, in the hope of understanding how spatial restriction affects stable distributions of wealth.

The assumptions about the space concept are as follows. There are regions and each player can stay in only one of the regions. Players can change their regions to pillage others. The regions are connected with one another, and thus players can travel from a region to another in one move. Players can form a coalition and combine their power only after getting together in a common region. If coalitions are in different regions, they cannot combine their power. The influence of the power of each coalition is limited within its region. Therefore, a coalition cannot pillage two other coalitions in different regions simultaneously.

The other assumptions of the spatial pillage game are the same as the original pillage games. A fixed amount of wealth is allocated among a finite number of players. Some players can form a coalition under the spatial restriction. A coalition can pillage less powerful coalitions within its region without any cost. An increase in the wealth of a coalition causes an increase in its power. Since the power of each coalition is endogenously determined, the spatial pillage game cannot have a characteristic function, which exogenously determines the power of each coalition.

The pillage games are characterized by *power functions* that determine the feasibility of pillages between coalitions. Jordan presented three power functions classified by the degree of their dependence on the sizes of coalitions. *Wealth is power* is one of the power functions and specifies the power of each coalition as its total wealth. Therefore, "wealth is power" is characterized as independent of the sizes of coalitions. Only the pillage game with this function has a stable set in every possible case. Therefore, the spatial pillage game adopts "wealth is power" as a power function.

As criteria for stable distributions of wealth and players, three solution concepts are explored, i.e., core, stable set, and farsighted core. First, the core is the collection of states at which pillage is not possible, thus it is one of the most persuasive solution concepts. However, because of its strong requirement, the core is too small to represent stable situations as shown in Theorem 13. Second, a stable set is much bigger than the core if it exists, as shown in Theorem 28. A stable set is a collection of states that is both internally stable and externally stable. Internal stability requires that pillage not be possible between states in the collection and external stability requires that pillage at a state outside the collection result in another state inside the collection. In most cases, however, there are no stable set. And even when they exist, they contain implausible states as shown in Theorems 28 and 35. Third, farsighted core, which was introduced by Jordan (2005), solves these problems in stable sets, as shown in Theorem 40 and Lemma 41. A farsighted core is a collection of states at which *pillage in expectation* is not possible in the sense that some members of the pillage would end up being worse off, and consequently they would not join the pillage.

In section 2.2, we search for the core and stable sets. First, the core is characterized. Then, since one-player models and two-player models are trivial, we start from three-player models and completely characterize stable sets in those models. Finally, we show that no stable set exists in a $I-$player and $N-$region model where $I = 4$ and $N = 2$ or $I \geq 4$ and $N \geq 3$. In section 2.3, we construct a consistent expectation, defined in Definition 38, to find a farsighted core. After confirming the existence of a farsighted core in a consistent expectation, we explore one of the properties that farsighted cores of consistent expectations have in common. Then, we show that in a $I-$player and $N-$region model where $1 \leq I \leq 3$ or $N = 1$, there is the unique farsighted core of consistent expectations. In section 2.4, a suggestion for further research is presented.

2.2. Core and stable set

The environment of a spatial pillage game is defined in Definitions 8 and 9. We normalize the total wealth to unity. Note that definitions in this section are applied throughout the whole paper.

Definition 8. [1]*The finite set* I *is the set of* ***players****. A* ***coalition*** *is a subset of* I*. The set* $A = \{w \in \mathbb{R}^I : w_i \geq 0$ *for all* $i \in I$ *and* $\sum_{z \in I} w_z = 1\}$ *is the set of* ***allocations****.*

The definitions below concern the spatial environment.

Definition 9. *The finite set* R *is the set of* ***regions*** *and the Cartesian product* R^I *is the set of* ***distributions****. Given a distribution* $p \in R^I$*, the coalition* $p^r = \{i \in I : p_i = r\}$ *is the* ***population*** *at region* r*.*

A distribution is short for a population distribution and denotes how players are distributed over the regions. For example, the distribution $p = (1, 1, 2)$ expresses that players 1 and 2 are at region 1 and player 3 is at region 2. Also, it means $p^1 = \{1, 2\}$ and $p^2 = \{3\}$.

A **state** denotes both the allocation and distribution of the status quo.

[1] *We follow notations in Jordan (2005).*

Definition 10. *The Cartesian product* $X = A \times R^I$ *is the set of* ***states****.*

For instance, the ordered pair $(w, p) = ((\frac{1}{2}, \frac{1}{4}, \frac{1}{4}), (1, 1, 2))$ is a state in the three-player and two-region model. The state (w, p) expresses that player 1 has $\frac{1}{2}$ and player 2 has $\frac{1}{4}$ while staying at region 1 and player 3 has $\frac{1}{4}$ while staying at region 2.

The **dominance** relation between states is defined as follows.

Definition 11. *Given states* (w, p) *and* (w', p')*, define* $W = \{i : w'_i > w_i\}$ *and* $L = \{i : w'_i < w_i\}$*. Suppose for some* $r, q \in R$*, i)* $\{i : w'_i \neq w_i\} \subset p'^r$*; ii)* $\{i : p_i \neq p'_i\} = \varnothing$ *or* $\{i : p_i \neq p'_i\} = W \subset p^q$*; and iii)* $\sum_{i \in W} w_i > \sum_{i \in L} w_i$*. Then* (w', p') ***dominates*** (w, p)*.*

The dominance relation shows which state the status quo can move to. It must satisfy both *physical* and *spatial conditions.* The physical condition requires that the winning coalition W must have enough power to pillage the losing coalition L. Definition 11 presents this condition at *iii*). Jordan (2005) introduced a variety of physical conditions. The condition *iii*) above accords with the physical condition of the *wealth is power* in Jordan (2005). The spatial condition requires that the act of pillaging must satisfy spatial restriction. This condition is expressed at *i*) and *ii*) in Definition 11. The condition *i*) means that transfers of wealth happen only in destination region r where the pillage happens. The condition $ii)$ denotes that only the winners can travel and that they are all from the common region q.

In this section, we adopt the solution concepts of **core** and **stable set**. The definition stated below follows Lucas (1992) and Jordan (2005).

Definition 12. *The set of undominated states is the* ***core*** *C. For any set E of states, let the set $U(E)$ be the set of states that are not dominated by any state in E. A set S of states is a* ***stable set*** *if it satisfies both $S \subset U(S)$, which means internal stability, and $S \supset U(S)$, which means external stability.*

Therefore, a stable set S is defined by the set of states that satisfies $S = U(S)$.

Theorem 13 embodies the core. Note that this result is applied throughout section 2.2.

Theorem 13. *The set $\{(w,p) \in X$: for each i, $w_i = \frac{1}{\#\{r \in R: \Sigma_{j \in p^r} w_j > 0\}}, \frac{1}{2}$, or $0\}$ is the core C.*

Proof. Suppose $(w,p) \in \{(w,p) \in X$: for each i, $w_i = \frac{1}{\#\{r \in R: \Sigma_{j \in p^r} w_j > 0\}}, \frac{1}{2}$, or $0\}$. If $w_i > 0$, then $w_i \geq \min\{\frac{1}{2}, \frac{1}{\#\{r \in R: \Sigma_{j \in p^r} w_j > 0\}}\}$. If $\frac{1}{\#\{r \in R: \Sigma_{j \in p^r} w_j > 0\}} \geq \frac{1}{2}$, then $\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\} = 1$ or 2 and thus for each i, $w_i = 1, \frac{1}{2}$, or 0. In this case, any coalition W cannot pillage another coalition L such that $W \cap L = \varnothing$ because if $\sum_{i \in L} w_i > 0$, then $\sum_{i \in L} w_i \geq \frac{1}{2}$ and so $\frac{1}{2} \geq \sum_{i \notin L} w_i \geq \sum_{i \in W} w_i$. If

$\frac{1}{2} > \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$, then $\#\{i : w_i > 0\} = \#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}$ since $\#\{i : w_i > 0\} \geq \#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}$ and $\#\{i : w_i > 0\} \times \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$ $= \#\{i : w_i > 0\} \times \min\{\frac{1}{2}, \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}\} \leq \Sigma_{i \in I} w_i = 1$, and thus for each i, $w_i = \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$ or 0 since $w_i \geq \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$ or 0. In this case, we have for each $r \in R$, $\Sigma_{j \in p^r} w_j = \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$ or 0, and so any coalition W such that $W \subset p^q$ for some $q \in R$ cannot pillage another coalition L such that $W \cap L = \varnothing$ because if $\sum_{i \in L} w_i > 0$, then $\sum_{i \in L} w_i \geq \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$ and $\frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}} \geq \sum_{i \in W} w_i$. Therefore, (w, p) is not dominated. Since (w, p) is arbitrary, every state in the set $\{(w, p) \in X :$ for each i, $w_i = \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}, \frac{1}{2}$, or $0\}$ is not dominated.

Suppose $(w, p) \notin \{(w, p) \in X :$ for each i, $w_i = \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}, \frac{1}{2}$, or $0\}$. Then there exists i such that $w_i \notin \{0, \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}, \frac{1}{2}, 1\}$. If $w_i > \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}$, then there exists $q \in R$ such that $\frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}} > \Sigma_{j \in p^q} w_j > 0$ since $\Sigma_{j \in I} w_j = 1$, and thus player i can pillage another player j such that $w_j > 0$ and $p_j = q$ since $w_i > \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}} > \Sigma_{k \in p^q} w_k \geq w_j > 0$. If $\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\} = 1$ and $w_i < 1$, then either $1 > w_i > \frac{1}{2}$ or $\frac{1}{2} > w_i > 0$ since $w_i \notin \{\frac{1}{2}, 0\}$, and thus player i can pillage player j such that $w_j > 0$ or the coalition $W = \{k : k \neq i$ and $p_k = p_i\}$ can pillage player i. If $\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\} \geq 2$ and $\frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}} > w_i$, then $\frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}}$, which denotes the average wealth of regions except the region p_i, is well defined, and thus either $w_i \geq \frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}}$ or

$\frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}} > w_i$. If $w_i \geq \frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}}$, then $\Sigma_{j \in p^{p_i}} w_j > \frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}}$, which means that the wealth of the region p_i is greater than the average wealth of regions except the region p_i, since $\frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}} > w_i \geq \frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}}$, and thus all players in the region p_i can pillage another region q such that $w_i \geq \Sigma_{j \in p^q} w_j > 0$. If $\frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}} > w_i$, then all players in q such that $\Sigma_{j \in p^q} w_j \geq \frac{\Sigma_{j \notin p^{p_i}} w_j}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0 \text{ and } r \neq p_i\}}$ can pillage the player i. This means that (w, p) is dominated by some state in X. Since (w, p) is arbitrary, every state in $X \setminus \{(w, p) \in X$: for each i, $w_i = \frac{1}{\#\{r \in R : \Sigma_{j \in p^r} w_j > 0\}}, \frac{1}{2}$, or $0\}$ is dominated. $\square$

2.2.1. Stable set in three-player models

To characterize stable sets, we divide states into four groups according to their distributions and allocations; group 1) all players are in one region; group 2) players have less than halves and occupy two regions; group 3) only one player has a half or more and the player stays alone in his region; and group 4) only one player has a half or more and the player is together with only another player in his region. It is easy to analyze the states in groups 1), 2), and 3) to find a stable set, it is not in group 4), however. Thus we would devote most this subsection to analyzing the states in group 4). For simplicity of expression, we call a state in group 4) a *basic state* and a set of basic states a *basic set*. Note that definitions and results in this subsection are applied to subsection 2.2.2 as well as subsection 2.2.1.

Definition 14 formalizes a **basic set** and a **basic state**.

Definition 14. *For any three distinct players i, j, and k, define the set $B(i; j, k)$ of distributions by $B(i; j, k) = \{p \in R^I$: for some region $r \in R$, $p^r = \{i, j\}$ or $\{i, k\}$ $\}$ and define the correspondence $B^i_{j,k} : [\frac{1}{2}, 1] \times R^I \longrightarrow\longrightarrow X$ by $B^i_{j,k}(a, \dot{p}) = \{(w, p) \in X : p = \dot{p},\ w_i \geq a,$ and $w_i + w_j + w_k = 1\}$. For each $p \in B(i; j, k)$, the set $B^i_{j,k}(\frac{1}{2}, p)$ of states is called a **basic set.** A state in a basic set is called a **basic state.***

The set $B(i; j, k)$ denotes the set of distributions such that either player i and player j, or player i and player k constitute all population in some region. For example, let $p = (1, 1, 2)$ and $p' = (1, 2, 1)$, then $p, p' \in B(1; 2, 3)$ because player 1 shares region 1 only with player 2 at the distribution p and only with player 3 at the distribution p'. The basic sets are visualized on the hyperplane of states in Figure 2.1. The black area and the gray area denote the basic set $B^1_{2,3}(\frac{1}{2}, (1, 1, 2))$ and the basic set $B^2_{1,3}(\frac{1}{2}, (1, 1, 2))$, respectively. They are all possible basic sets under the distribution $(1, 1, 2)$.

In Figure 2.1, consider the basic state $(w, p) = ((\frac{7}{12}, \frac{3}{12}, \frac{1}{6}), (1, 1, 2))$ where players 1 has $\frac{7}{12}$ and player 2 has $\frac{3}{12}$ while staying at region 1 and player 3 has $\frac{1}{6}$ while

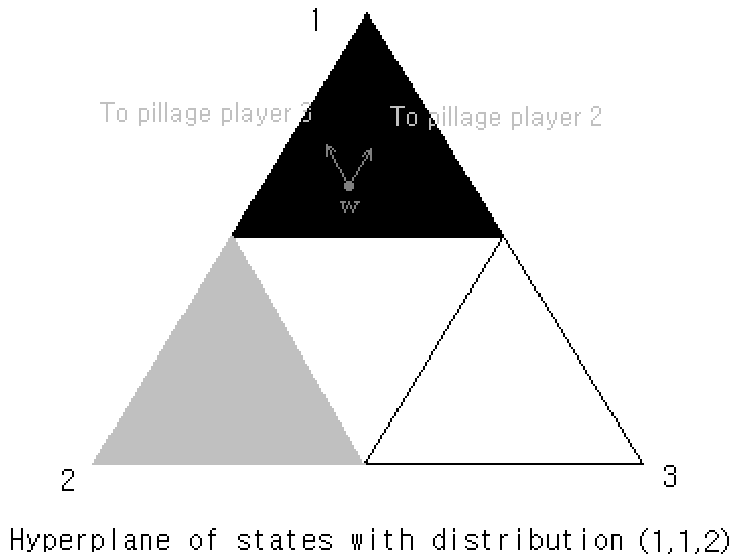

Figure 2.1. Basic Sets

staying at region 2. Note that player 1 cannot pillage players 2 and 3 simultaneously because players 2 and 3 are in different regions. If player 1 pillages player 3 at (w, p), then the allocation of the state is located on the left arrow in the figure, and the distribution changes from $(1, 1, 2)$ to $(2, 1, 2)$. If player 1 pillages player 2 at (w, p), then the state is located on the right arrow, and the distribution does not change.

For notational simplicity, we define the following set of states.

Definition 15. *For any three distinct players i, j, and k, define the correspondence $H^i_{j,k} : [\frac{1}{2}, 1] \times R^I \longrightarrow\!\!\!\!\longrightarrow X$ by $H^i_{j,k}(a, \dot{p}) = \{(w, p) \in X : p = \dot{p},\ w_i = a,$ and $w_i + w_j + w_k = 1\}$.*

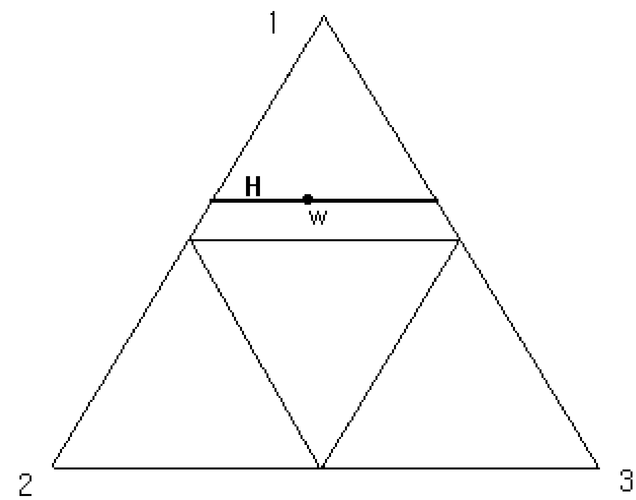

Figure 2.2. The Set $H^1_{2,3}(\frac{7}{12},(1,1,2))$

For each $(a,p) \in [\frac{1}{2},1] \times R^I$, the set $H^i_{j,k}(a,p)$ consists of the states such that $w_i = a$ in $B^i_{j,k}(\frac{1}{2},p)$. In Figure 2.2, the bold horizontal line and the dot denote $H^i_{j,k}(\frac{7}{12},(1,1,2))$ and $w = (\frac{7}{12},\frac{3}{12},\frac{1}{6})$, respectively.

Definition 16 introduces the condition that a stable set has to satisfy. The condition is related to basic sets, and thus we call this condition the **basic condition**. If a set S' of states lacks the basic condition, then S' cannot satisfy internal stability and external stability simultaneously.

Definition 16. *Given a set E of states, for any two distinct states $(w,p),(\dot{w},p) \in E \cap B^i_{j,k}(\frac{1}{2},p)$ such that $p \in B(i;j,k)$ and $1 > \dot{w}_i \geq w_i > \frac{1}{2}$, suppose that i)*

$0 < \dot{w}_j \leq w_j$ and $0 < \dot{w}_k \leq w_k$; and ii) $\dot{w}_k < w_k$ when $p_i = p_j$ and $\dot{w}_j < w_j$ when $p_i = p_k$. Then the set E of states is said to satisfy the ***basic condition****.*

We can prove that a stable set satisfies the basic condition by way of contradiction. That is, if we assume that there is a stable set that lacks the basic condition, then we can show that the stable set cannot satisfy external stability and internal stability simultaneously.

Lemma 17. *A stable set satisfies the basic condition.*

Proof. By way of contradiction, suppose that there exists a stable set S that does not satisfies the basic condition. Then for some three distinct players i, j, and k, there exist two distinct states $(w, p), (\dot{w}, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ such that $p \in B(i; j, k)$; $1 > \dot{w}_i \geq w_i > \frac{1}{2}$; if $p_i = p_k$, then $\dot{w}_k > w_k$ or $\dot{w}_j \geq w_j$; and if $p_i = p_j$, then $\dot{w}_k \geq w_k$ or $\dot{w}_j > w_j$. Without loss of generality, we can assume that player i is together with player j in a common region, i.e., $p_i = p_j$. Then we must have that either $\dot{w}_k \geq w_k$ or $\dot{w}_j > w_j$. We can show that in each case, S cannot satisfy internal stability and external stability simultaneously.

Suppose that we have that $\dot{w}_k \geq w_k$. We first show that $\dot{w}_k > w_k$. Every state $(\ddot{w}, p)$ in $B^i_{j,k}(\frac{1}{2}, p)$ such that $\ddot{w}_k = w_k$ and $\ddot{w}_i = w_i$ has $\ddot{w}_j = w_j$ since $\ddot{w}_i + \ddot{w}_j + \ddot{w}_k = w_i + w_j + w_k = 1$. Thus we have that $(\ddot{w}, p) = (w, p)$. Therefore, $(\dot{w}, p)$ cannot have $\dot{w}_k = w_k$ and $\dot{w}_i = w_i$ since $(w, p) \neq (\dot{w}, p)$. Every state $(\dddot{w}, p)$

in $B^i_{j,k}(\frac{1}{2}, p)$ such that $\ddot{w}_k = w_k$ and $\ddot{w}_i > w_i$ is the state that results from player i pillaging player j at the state (w, p); that is, such state $(\ddot{w}, p)$ dominates (w, p). By internal stability, S cannot contain such state $(\ddot{w}, p)$ and thus $(\dot{w}, p)$ cannot be $\dot{w}_k = w_k$ and $\dot{w}_i > w_i$. Therefore, we must have that $\dot{w}_k \neq w_k$ and thus that $\dot{w}_k > w_k$.

Let the allocation w' be $w'_j = \dot{w}_j$, $w'_k = w_k$, and $w'_i = 1 - \dot{w}_j - w_k$. Since $\dot{w}_i \geq w_i$ and $\dot{w}_k > w_k$, we have $w_j = 1 - w_i - w_k > 1 - \dot{w}_i - \dot{w}_k = \dot{w}_j$. Thus we have $w'_i = 1 - \dot{w}_j - w_k > 1 - w_j - w_k = w_i$. Since $w'_k = w_k$, $w'_i > w_i$, and $w'_j = \dot{w}_j = 1 - w_k - w'_i = w_j - (w'_i - w_i)$, (w', p) dominates (w, p) by player i pillaging player j. Thus S cannot contain (w', p) according to internal stability. To satisfy external stability, S has to dominate (w', p).

However, we can show that S cannot dominate (w', p). The stable set S can dominate (w', p), only if S contains those states as follows; the states that result from player i pillaging player j at (w', p), the states that result from player i pillaging player k at (w', p), the states that result from players i and j pillaging player k at (w', p), the states that result from player j pillaging player k at (w', p) when $w'_j > w'_k$, or the states that result from player k pillaging player j at (w', p) when $w'_k > w'_j$. Note that player j and player k are in different regions and so player i cannot pillage both of them simultaneously although player i has enough power to do it, i.e., $p_j \neq p_k$ and $w'_i > w'_j + w'_k$. We will show that S cannot contain any state above.

Every state that results from player i pillaging player j at (w', p) dominates (w, p), which is in S according to our assumption. By internal stability, S cannot contain those states. Every state that results from player i pillaging player k at (w', p) dominates $(\dot{w}, p)$, which is in S according to our assumption. Similarly, S cannot contain those states. The states that result from players i and j pillaging player k at (w', p) are all dominated by $((0, ..., w_i = 1, ..., 0), (p_k, ..., p_k))$, which is in the core. Thus S cannot contain those states. The states that result from player j pillaging player k at (w', p) or that result from player k pillaging player j at (w', p) are all dominated by either $((0, ..., w_i = 1, ..., 0), (p_k, ..., p_k))$ or $((0, ..., w_i = 1, ..., 0), (p_j, ..., p_j))$. Thus S cannot contain those states. Therefore, S cannot dominate (w', p) and thus cannot satisfy external stability. This contradiction shows that $\dot{w}_k \geq w_k$ is not possible.

Suppose that we have that $\dot{w}_j > w_j$. Then we can similarly show that S cannot dominate the state (w'', p) such that $w''_j = w_j$, $w''_k = \dot{w}_k$, and $w''_i = 1 - w_j - \dot{w}_k$. Consequently, the stable set S cannot satisfy internal stability and external stability simultaneously. This contradiction completes the proof. □

Lemma 18 presents another condition that a stable set must follow. Lemma 17 examines the relation between two basic states in a stable set. Lemma 18 examines the relation between a basic state and another state whose distribution results from the move of the player who has a half or more at the basic state.

Lemma 18. *Suppose that* $p \in B(i; j, k)$ *and* $(w, p) \in B^i_{j,k}(\frac{1}{2}, p)$*. Let a distribution* $\dot{p}$ *satisfy that* $\dot{p}_z = p_z$ *for each* $z \neq i$ *and* $\dot{p}_i \in \{p_j, p_k\}$*. Given a stable set* S*, if* $(w, \dot{p}) \in S$*, then* $(w, p) \in S$*.*

Proof. If $\dot{p}_i = p_i \in \{p_j, p_k\}$, then $\dot{p} = p$, and thus this result obviously follows. Now, we have to show that if $\dot{p}_i \neq p_i$ and $(w, \dot{p}) \in S$, then $(w, p) \in S$. Suppose by way of contradiction that $\dot{p}_i \neq p_i$ and $(w, \dot{p}) \in S$, but $(w, p) \notin S$. It suffices to show that S cannot contain any state that dominates (w, p).

Without loss of generality, we assume that $w_j \geq w_k$. Since $(w, \dot{p}) \in B^i_{j,k}(\frac{1}{2}, \dot{p})$, we have that $w_i \geq \frac{1}{2}$ and $w_i + w_j + w_k = 1$. We first show that if $w_j > w_k$ then $w_j + w_k < \frac{1}{2}$. By way of contradiction, suppose not, that is, $w_j > w_k$ and $w_j + w_k = \frac{1}{2}$. Then $(w, \dot{p})$ is dominated by the state $(\ddot{w}, \ddot{p})$ such that $\ddot{w}_i = \ddot{w}_j = \frac{1}{2}$, which is in the core, C, by player j pillaging all wealth of player k at $(w, \dot{p})$. This contradicts internal stability of S since $(w, \dot{p}) \in S$ and $C \subset S$.

Let (w', p') result from player j or players i and j pillaging player k at (w, p). Then we have $p'_j = p'_k$. If players i and j pillage player k at $(w, \dot{p})$, then $w'_i > w_i \geq \frac{1}{2}$

and $w'_j > 0$, and thus player i can deprive the other players of their all wealth by pillage since $p'_j = p'_k$ and $w'_j + w'_k < w'_i$. If player j alone pillages player k at $(w, \dot{p})$, then $w_j > w_k$ and thus $w'_i = w_i > \frac{1}{2}$ since $w_j + w_k < \frac{1}{2}$. Thus player i can also deprive the other players of their all wealth in one move since $p'_j = p'_k$ and $w'_j + w'_k < w'_i$. Therefore, (w', p') is dominated by some state $(\dot{w}', \dot{p}')$ in the core such that $\dot{w}'_i = 1$, and thus S cannot contain (w', p'). Similarly, we can show that S cannot contain any state that results from players i and k pillaging player j at (w, p).

Let (w'', p'') result from player i pillaging player j at (w, p). Then we have that $w''_i > w_i$, $w''_j < w_j$, and $w''_z = w_z$ for each $z \in I \setminus \{i, j\}$. Note that $\{z : p''_z \neq p_z\} \subset \{i\}$ and thus $\{z : p''_z \neq \dot{p}_z\} \subset \{i\}$ since $\dot{p}_z = p_z$ for each $z \neq i$. Therefore, (w'', p'') dominates $(w, \dot{p})$ by player i pillaging player j. Thus S cannot contain (w'', p''). Similarly, we can show that S cannot contain any state that results from player i pillaging player k at (w, p).

Consequently, S cannot contain any state that dominates $(\dot{w}, p)$ and thus cannot satisfy external stability. This contradiction completes the proof. $\square$

Lemma 19 shows another implication of Lemma 17 and Lemma 18. We will express basic states in a stable set with a function. Lemma 19 provides a basis to define the function that characterizes a stable set.

Lemma 19. *Given a stable set S, $S \cap H^i_{j,k}(a,p)$ has a single element for each $1 \geq a > \frac{1}{2}$ and $p \in B(i;j,k)$.*

Proof. It suffices to show that for each $1 > a > \frac{1}{2}$ and $p \in B(i;j,k)$, $S \cap H^i_{j,k}(a,p)$ has a single element because $H^i_{j,k}(1,p)$ has only one state regardless of p, which is in the core and so in a stable set. Suppose that $(w',p),(w,p) \in S \cap H^i_{j,k}(a,p)$ such that $1 > a > \frac{1}{2}$ and $p \in B(i;j,k)$. Then we have that $(w',p),(w,p) \in S \cap B^i_{j,k}(\frac{1}{2},p)$ and $1 > w'_i = w_i > \frac{1}{2}$. Suppose by way of contradiction that $w' \neq w$. By the basic condition of S, we have that either $w'_j \leq w_j$ and $w'_k < w_k$, or $w'_j < w_j$ and $w'_k \leq w_k$ since $1 > w'_i \geq w_i > \frac{1}{2}$. However, neither case is possible since $w'_i + w'_j + w'_k = w_i + w_j + w_k$. Therefore, we must have that $w' = w$, and thus $S \cap H^i_{j,k}(a,p)$ has at most one state for each $1 > a > \frac{1}{2}$ and $p \in B(i;j,k)$.

We need to show that $S \cap H^i_{j,k}(a,p) \neq \varnothing$ for each $1 > a > \frac{1}{2}$ and $p \in B(i;j,k)$ to complete the proof. By way of contradiction, suppose that there exists a stable set S such that $S \cap H^i_{j,k}(a,p) = \varnothing$ for some $1 > a > \frac{1}{2}$ and $p \in B(i;j,k)$. Without loss of generality, we can assume that player i is together with player j in a common region, i.e., $p_i = p_j$. Let $\bar{w}_j = \sup\{w_j : (w,p) \in S \cap B^i_{j,k}(a,p)\}$ and $\bar{w}_k = \sup\{w_k : (w,p) \in S \cap B^i_{j,k}(a,p)\}$.

We first show that $a + \bar{w}_j + \bar{w}_k \leq 1$. Suppose by way of contradiction that $a + \bar{w}_j + \bar{w}_k > 1$. Then by the definitions of $\bar{w}_j$ and $\bar{w}_k$, there exist states $(\dot{w},p)$

and $(\ddot{w}, p)$ such that $(\dot{w}, p), (\ddot{w}, p) \in S \cap B^i_{j,k}(a, p)$ and $a + \dot{w}_j + \ddot{w}_k > 1$. Since $(\dot{w}, p), (\ddot{w}, p) \in B^i_{j,k}(a, p)$, we have that $\dot{w}_i, \ddot{w}_i \geq a$ and thus that $\dot{w}_i + \dot{w}_j + \ddot{w}_k > 1$ and $\ddot{w}_i + \dot{w}_j + \ddot{w}_k > 1$. Thus we have that $\dot{w}_j > \ddot{w}_j$ and $\dot{w}_k < \ddot{w}_k$ since $\dot{w}_i + \dot{w}_j + \dot{w}_k = \ddot{w}_i + \ddot{w}_j + \ddot{w}_k = 1$. However, the basic condition of S implies that both $\dot{w}_j \geq \ddot{w}_j$ and $\dot{w}_k \geq \ddot{w}_k$ if $\dot{w}_i \leq \ddot{w}_i$ and both $\dot{w}_j \leq \ddot{w}_j$ and $\dot{w}_k \leq \ddot{w}_k$ if $\dot{w}_i \geq \ddot{w}_i$. This contradiction guarantees that $a + \bar{w}_j + \bar{w}_k \leq 1$.

Define the allocation w to be $w_i = a$, $w_j = \bar{w}_j + \frac{1-(a+\bar{w}_j+\bar{w}_k)}{2}$, and $w_k = \bar{w}_k + \frac{1-(a+\bar{w}_j+\bar{w}_k)}{2}$. Then S cannot contain (w, p) since $(w, p) \in H^i_{j,k}(a, p)$ and $S \cap H^i_{j,k}(a, p) = \varnothing$. To prove that the assertion, $S \cap H^i_{j,k}(a, p) = \varnothing$, is impossible, it suffices to show that S cannot dominate (w, p).

First, we show that every state that results from player i pillaging player j at (w, p) cannot be in S. Let the state (w', p) result from player i pillaging player j at (w, p). Then we have that $w'_i > a$, $w'_j < w_j$, $w'_k = w_k$, and $(w', p) \in B^i_{j,k}(a, p)$; that is, player i increases its wealth through pillaging player j at the state (w, p) and player k maintains its wealth because the pillage does not affect player $k's$ wealth. If $1 > (\bar{w}_j + \bar{w}_k + a)$ then $w'_k > \bar{w}_k$, and thus $(w', p) \notin S$ because $\bar{w}_k$ is the supremum of the wealth that player k can have at states in $S \cap B^i_{j,k}(a, p)$ and $(w', p) \in B^i_{j,k}(a, p)$. If $1 = \bar{w}_j + \bar{w}_k + a$, then $w'_j < w_j = \bar{w}_j$ and $w'_k = w_k = \bar{w}_k$. Thus there exists a state $(w'', p) \in S \cap B^i_{j,k}(a, p)$ such that $w'_j < w''_j \leq \bar{w}_j$ and $w''_k \leq \bar{w}_k$ by the definitions of $\bar{w}_j$ and $\bar{w}_k$. Thus if $(w', p) \in S$, then the basic condition of S

means that $w''_i < w'_i$ since $w'_j < w''_j$ and so that $w'_k < w''_k$. Since $w''_k \leq \bar{w}_k = w'_k$, we have that $(w', p) \notin S$. Note that (w', p) is arbitrary such that (w', p) results from player i pillaging player j at (w, p). Therefore, S cannot contain the states that result from player i pillaging player j at (w, p).

Second, we prove that every state that results from player i pillaging player k at (w, p) cannot be in S. Suppose by way of contradiction that S contains a state (w''', p') that results from player i pillaging player k at (w, p). Then we have that $w'''_i > w_i$, $w'''_k < w_k$, and $w'''_j = w_j$. Consider the state (w''', p). Then we have that $p \in B(i; j, k)$ and $p'_z = p_z$ for each $z \neq i$ and $p'_i = p_k$. Lemma 18 means that $(w''', p) \in S \cap B^i_{j,k}(a, p)$ since $(w''', p') \in S$. Then we have that $\bar{w}_j \geq w'''_j$ according to the definition of $\bar{w}_j$. Since $w'''_j = w_j = \bar{w}_j + \frac{1-(a+\bar{w}_j+\bar{w}_k)}{2} \geq \bar{w}_j$, we have that $w'''_j = \bar{w}_j$ and thus that $1 = a + \bar{w}_j + \bar{w}_k$. By the definition of $\bar{w}_k$, there exists $(w^{(4)}, p)$ such that $(w^{(4)}, p) \in S \cap B^i_{j,k}(a, p)$ and $w'''_k < w^{(4)}_k < w_k = \bar{w}_k$. The basic condition of S implies that $w'''_j \leq w^{(4)}_j$ since $w'''_k < w^{(4)}_k$. Since $w^{(4)}_j \leq \bar{w}_j = w'''_j$ according to the definition of $\bar{w}_j$, we have that $w'''_j = w^{(4)}_j$. Therefore, we have that $w'''_i > w^{(4)}_i$, $w'''_k < w^{(4)}_k$, and $w'''_j = w^{(4)}_j$. This means that (w''', p') dominates $(w^{(4)}, p)$ by player i pillaging player k at $(w^{(4)}, p)$. This contradiction assures that $(w''', p') \notin S$. Since (w''', p') is arbitrary, S cannot contain the states that result from player i pillaging player k at (w, p).

Finally, we demonstrate that every state that dominates (w, p) and that is not covered by the two cases above is not in S. Note that these states result from either player j or player k moving to the other regardless of the move of player i. Consequently, player j and player k are in a common region at these states. Therefore, all such states are dominated by some state in the core such that player i has all of the wealth because player i, who has a majority of the power, $w_i \geq a > \frac{1}{2}$, can pillage both players in one move. Therefore, S cannot contain these states.

Consequently, S cannot dominate (w, p). This means that S cannot satisfy internal stability and external stability simultaneously. This contradiction guarantees that we must have that $S \cap H^i_{j,k}(a, p) \neq \varnothing$ for each $1 > a > \frac{1}{2}$ and $p \in B(i; j, k)$. □

Definition 20 presents the conditions for the function that characterizes a stable set and names the function a **basic function**.

Definition 20. *For any three distinct players* $i, j,$ *and* k*, let a function* $\beta^i_{j,k}$: $[0, \frac{1}{2}] \times B(i; j, k) \longrightarrow [0, \frac{1}{2}]$ *satisfy that* $\beta^i_{j,k}(\lambda, p) \leq \lambda$ *for each* $(\lambda, p) \in [0, \frac{1}{2}] \times B(i; j, k)$*. Define the set* $B(\beta^i_{j,k})$ *of states by* $B(\beta^i_{j,k}) = \{(w, p) \in \bigcup_{p \in B(i;j,k)} B^i_{j,k}(\frac{1}{2}, p) :$ *for some* $(\lambda, p) \in [0, \frac{1}{2}] \times B(i; j, k)$*,* $w_j = \beta^i_{j,k}(\lambda, p)$ *and* $w_k = \lambda - \beta^i_{j,k}(\lambda, p)\}$*. Suppose that* $\beta^i_{j,k}$ *satisfies three conditions as follows; i)* $B(\beta^i_{j,k})$ *satisfies the basic* *condition; ii) if $p, \dot{p} \in B(i; j, k)$ satisfy that $\dot{p}_z = p_z$ for each $z \neq i$ and $p_i \neq \dot{p}_i$, then for each $\lambda \in [0, \frac{1}{2}]$, $\beta^i_{j,k}(\lambda, p) = \beta^i_{j,k}(\lambda, \dot{p})$; and for each $p \in B(i; j, k)$, iii) if $\lim_{\lambda \to 1/2} \beta^i_{j,k}(\lambda, p) = \frac{1}{4}$, then $\beta^i_{j,k}(\frac{1}{2}, p) = \frac{1}{4}$, otherwise $\beta^i_{j,k}(\frac{1}{2}, p) = \frac{1}{2}$. Then $\beta^i_{j,k}$ is called a* ***basic function****.*

Lemma 21 characterizes the functions that generate the set satisfying the basic condition.

Lemma 21. *Let a function $\beta^i_{j,k}$ be a function from $[0, \frac{1}{2}] \times B(i; j, k)$ to $[0, \frac{1}{2}]$ such that $\beta^i_{j,k}(\lambda, p) \leq \lambda$ for each $(\lambda, p) \in [0, \frac{1}{2}] \times B(i; j, k)$. If $B(\beta^i_{j,k})$ satisfies the basic condition, then $\beta^i_{j,k}(\cdot, p)$ is uniformly continuous and non-decreasing on $[0, \frac{1}{2})$.*

Proof. If $B(\beta^i_{j,k})$ satisfies the basic condition, then for each $\frac{1}{2} > \lambda > \lambda' \geq 0$ and $p \in B(i; j, k)$, we have that $\lambda - \beta^i_{j,k}(\lambda, p) \geq \lambda' - \beta^i_{j,k}(\lambda', p)$ and $\beta^i_{j,k}(\lambda, p) \geq \beta^i_{j,k}(\lambda', p)$. Therefore, Given any $\varepsilon > 0$, we must have that $\varepsilon > \beta^i_{j,k}(\lambda, p) - \beta^i_{j,k}(\lambda', p) \geq 0$ for all $\lambda, \lambda' \in [0, \frac{1}{2})$ and $p \in B(i; j, k)$ such that $\varepsilon > \lambda - \lambda' \geq 0$. This shows that the function $\beta^i_{j,k}(\cdot, p)$ is uniformly continuous and non-decreasing on $[0, \frac{1}{2})$. □

Corollary 22 shows properties of a basic function.

Corollary 22. *For each $p \in B(i; j, k)$, a basic function $\beta^i_{j,k}(\cdot, p) : [0, \frac{1}{2}] \to [0, \frac{1}{2}]$ is uniformly continuous and non-decreasing on $[0, \frac{1}{2})$.*

Proof. According to Lemma 21, this result follows. □

Lemma 23 strengthens Lemma 17. More concretely, Lemma 23 shows that a stable set must satisfy three conditions that are reflected on a basic function.

Lemma 23. *Given a stable set S, for any three distinct players i, j, and k, there exists a unique basic function $\beta^i_{j,k}$ such that $(B(\beta^i_{j,k}) \cup C) \cap B^i_{j,k}(\frac{1}{2}, p) = S \cap B^i_{j,k}(\frac{1}{2}, p)$ for each $p \in B(i; j, k)$.*

Proof. According to Lemma 19, $S \cap H^i_{j,k}(a, p)$ has a single state for each $1 \geq a > \frac{1}{2}$ and $p \in B(i; j, k)$. In addition, we have that $S \cap H^i_{j,k}(\frac{1}{2}, p) \neq \varnothing$ for each $p \in B(i; j, k)$ since such a set contains some states in C, at which two players have halves. Therefore, we can define the function $\alpha : [\frac{1}{2}, 1] \times B(i; j, k) \to [0, \frac{1}{2}]$ as follows; *i)* $\alpha(w_i, p) = w_j$ such that $(w, p) \in S \cap \bigcup_{p \in B(i;j,k)} B^i_{j,k}(\frac{1}{2}, p)$; and for each $p \in B(i; j, k)$, *ii)* if there exists $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ such that $w_i = \frac{1}{2}$ and $w_j = \frac{1}{4}$, then $\alpha(\frac{1}{2}, p) = \frac{1}{4}$, otherwise $\alpha(\frac{1}{2}, p) = \frac{1}{2}$. That is, the function α assigns each (w_i, p) the player $j's$ allocation according to $(w, p) \in S \cap \bigcup_{p \in B(i;j,k)} B^i_{j,k}(\frac{1}{2}, p)$.

Define the function $\beta^i_{j,k} : [0, \frac{1}{2}] \times B(i; j, k) \to [0, \frac{1}{2}]$ by $\beta^i_{j,k}(\lambda, p) = \alpha(1 - \lambda, p)$. Then it is easily seen that for each $(\lambda, p) \in [0, \frac{1}{2}] \times B(i; j, k)$, $\beta^i_{j,k}(\lambda, p) \leq \lambda$ and $((B(\beta^i_{j,k}) \cup C) \cap B^i_{j,k}(\frac{1}{2}, p)) \setminus H^i_{j,k}(\frac{1}{2}, p) = (S \cap B^i_{j,k}(\frac{1}{2}, p)) \setminus H^i_{j,k}(\frac{1}{2}, p)$. Next, we will show that for each $p \in B(i; j, k)$, $(B(\beta^i_{j,k}) \cup C) \cap H^i_{j,k}(\frac{1}{2}, p) = S \cap H^i_{j,k}(\frac{1}{2}, p)$.

By the definition of $\beta^i_{j,k}$, we have that $B(\beta^i_{j,k}) \subset S$ and thus that $(B(\beta^i_{j,k}) \cup C) \cap H^i_{j,k}(\frac{1}{2}, p) \subset S \cap H^i_{j,k}(\frac{1}{2}, p)$ for each $p \in B(i; j, k)$. Note that if $(w', p) \in H^i_{j,k}(\frac{1}{2}, p)$ with $w'_j \notin \{0, \frac{1}{4}, \frac{1}{2}\}$, then (w', p) is dominated by some state in C such that two players have halves. Therefore, if $(w, p) \in S \cap H^i_{j,k}(\frac{1}{2}, p)$ for some $p \in B(i; j, k)$, then $w_j = 0$, $\frac{1}{4}$, or $\frac{1}{2}$. And thus $(w, p) \in (B(\beta^i_{j,k}) \cup C) \cap H^i_{j,k}(\frac{1}{2}, p)$ because if $w_j = 0$ or $\frac{1}{2}$ then $(w, p) \in C$ and if $w_j = \frac{1}{4}$ then $(w, p) \in B(\beta^i_{j,k})$.

To complete the proof, we must show that the function $\beta^i_{j,k}$ is a basic function. Since $(B(\beta^i_{j,k}) \cup C) \cap B^i_{j,k}(\frac{1}{2}, p) = S \cap B^i_{j,k}(\frac{1}{2}, p)$ for each $p \in B(i; j, k)$, the set $B(\beta^i_{j,k})$ satisfies the basic condition as S does by Lemma 17. If $p, \dot{p} \in B(i; j, k)$ such that $\dot{p}_z = p_z$ for each $z \neq i$ and $p_i \neq \dot{p}_i$, then $S \cap H^i_{j,k}(w_i, p) = S \cap H^i_{j,k}(w_i, \dot{p})$ for each $1 \geq w_i \geq \frac{1}{2}$ by Lemma 18, and thus $\beta^i_{j,k}(\lambda, p) = \beta^i_{j,k}(\lambda, \dot{p})$ for each $\lambda \in [0, \frac{1}{2}]$. Now, we only need to prove that $\lim_{\lambda \to 1/2} \beta^i_{j,k}(\lambda, p) = \frac{1}{4}$ if and only if $\beta^i_{j,k}(\frac{1}{2}, p) = \frac{1}{4}$ because if $\beta^i_{j,k}(\frac{1}{2}, p) \neq \frac{1}{4}$, then $S \cap H^i_{j,k}(\frac{1}{2}, p)$ has two elements at which player j has either 0 or $\frac{1}{2}$.

First, we prove that for some $p \in B(i; j, k)$, if $\beta^i_{j,k}(\frac{1}{2}, p) = \frac{1}{4}$ then $\lim_{w_i \longrightarrow 1/2} \beta^i_{j,k}(w_i, p) = \frac{1}{4}$. Suppose that for some $p \in B(i; j, k)$, there exists $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ with $w_i = \frac{1}{2}$ and $w_j = \frac{1}{4}$. Without loss of generality, we assume that player i is together with player j in a common region, i.e., $p_i = p_j$. Since $\beta^i_{j,k}(\cdot, p)$ is uniformly continuous on $[0, \frac{1}{2})$ by Lemma 21, $\lim_{w_i \longrightarrow 1/2} \beta^i_{j,k}(w_i, p)$ always exists. Let $b = \lim_{w_i \longrightarrow 1/2} \beta^i_{j,k}(w_i, p)$. Suppose by way of contradiction that $b \neq \frac{1}{4}$.

Let $b > \frac{1}{4}$ first. Then there exists $\dot{w}_i$ such that $\beta^i_{j,k}(1 - \dot{w}_i, p) = \frac{1}{4}$ by the continuity of $\beta^i_{j,k}(\cdot, p)$ on $[0, \frac{1}{2})$, and thus there exists the allocation $\dot{w}$ such that $\dot{w}_j = \frac{1}{4}$ and $(\dot{w}, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$. Let $\dot{p} \in R^I$ such that $\dot{p}_z = p_z$ for each $z \neq i$ and $\dot{p}_i = p_k$. Then the state $(\dot{w}, \dot{p})$ dominates (w, p) by player i pillaging player k at (w, p) since $\dot{w}_i > w_i$, $\dot{w}_k < w_k$, and $\dot{w}_j = w_j$. Thus we have that $(\dot{w}, \dot{p}) \notin S$. However, every state that results from player i pillaging either player j or player k at $(\dot{w}, \dot{p})$ dominates $(\dot{w}, p) \in S$ as well. Every state that results from either player j or player k moving his region to pillage, regardless of the movement of player i, is dominated by some state in the core such that player i has the entire wealth. Therefore, S cannot dominate $(\dot{w}, \dot{p})$, and thus S lacks external stability.

Let $b < \frac{1}{4}$ next. Then there exists $\lambda \in [0, \frac{1}{2})$ with $\lambda - \beta^i_{j,k}(\lambda, p) > \frac{1}{4}$, and thus there exists λ'' such that $\lambda'' - \beta^i_{j,k}(\lambda'', p) = \frac{1}{4}$ since $\beta^i_{j,k}(\cdot, p)$ is continuous on $[0, \frac{1}{2})$. Then there exists $(\ddot{w}, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ with $\ddot{w}_k = \frac{1}{4}$ and $\ddot{w}_j = \beta^i_{j,k}(\lambda'', p)$, and $(\ddot{w}, p)$ dominates $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ by player i pillaging player j at (w, p). This shows that S lacks internal stability. Consequently, these contradictions ensure that $b = \frac{1}{4}$.

Lastly, we prove that for some $p \in B(i; j, k)$, if $\lim_{\lambda \longrightarrow 1/2} \beta^i_{j,k}(\lambda, p) = \frac{1}{4}$, then $\beta^i_{j,k}(\frac{1}{2}, p) = \frac{1}{4}$. Suppose that for some $p \in B(i; j, k)$, $\lim_{\lambda \longrightarrow 1/2} \beta^i_{j,k}(\lambda, p) = \frac{1}{4}$. Note that $(w, p) \in B^i_{j,k}(\frac{1}{2}, p)$ with $w_i = \frac{1}{2}$, $w_j = \frac{1}{4}$, and $p_i = p_j$ is dominated only either by player i pillaging another player at (w, p), or by players i and j pillaging player k at (w, p). Note that $\beta^i_{j,k}(\lambda, p)$ and $\lambda - \beta^i_{j,k}(\lambda, p)$ denote player $j's$ allocation and player $k's$ allocation, respectively, when player i has $1-\lambda$ at the distribution p in the stable set S. Because $\lim_{\lambda \longrightarrow 1/2} \beta^i_{j,k}(\lambda, p) = \frac{1}{4}$, we have that $\lim_{\lambda \longrightarrow 1/2} (\lambda - \beta^i_{j,k}(\lambda, p)) = \frac{1}{4}$. Therefore, the basic condition implies that a state (w'', p) such that $w''_j < \frac{1}{4}$ and $w''_k = \frac{1}{4}$ is not in S. Such a state (w'', p) results from player i pillaging player j at (w, p). Furthermore, the basic condition implies that a state (w''', p) such that $w'''_j = \frac{1}{4}$ and $w'''_k < \frac{1}{4}$ is not in S. By Lemma 18, S cannot contain such a state $(w''', \dot{p})$ that results from player i pillaging player k at (w, p). Finally, every state that results from players i and j pillaging player k at (w, p) is dominated by some state in the core such that player i has all of the wealth. Therefore, S must contain (w, p) to satisfy external stability, and thus we have that $\beta^i_{j,k}(\frac{1}{2}, p) = \frac{1}{4}$. $\square$

Jordan (2005) studied the pillage game of "wealth is power" power function without spatial restriction and found the unique stable set, the set of dyadic allocations. Definition 24 introduces a **dyadic state** that satisfies three conditions below and the set of dyadic states D. Theorem 25 establishes that the set D is the unique stable set in a one-region model. Note that Definition 24 and Theorem 25 are adapted from Jordan (2005) for the spatial pillage game.

Definition 24. *An allocation $w \in A$ is dyadic if for each i, $w_i = 0$ or $(\frac{1}{2})^{k_i}$ for some nonnegative integer k_i. A state (w, p) is **dyadic** if it satisfies that i) w is dyadic; ii) for each $r \in R$, $\sum_{i \in p^r} w_i = 0, \frac{1}{2}$, or 1; and iii) if there exists a region $r' \in R$ with $\sum_{i \in p^{r'}} w_i = \frac{1}{2}$, then there exists player z with $w_z = \frac{1}{2}$. The set D denotes the set of dyadic states.*

Theorem 25 (Theorem 3.3 in Jordan, 1999)**.** *In a one-region model, the unique stable set is D.*

Lemma 26 reveals another implication of Theorem 25. It applies Theorem 25 to a general model, which possibly can have more than one region.

Lemma 26. *Define the set $\bar{X}$ of states by $\bar{X} = \{(w, p) \in X :$ for each region $r \in R$, $\sum_{i \in p^r} w_i = 0, \frac{1}{2}$, or 1 and if for some region $r' \in R$, $\sum_{i \in p^{r'}} w_i = \frac{1}{2}$, then for some player z, $w_z = \frac{1}{2}\}$. Then D is the unique set that satisfies both internal stability and external stability with respect to $\bar{X}$. In addition, a stable set includes D.*

Proof. For any region $r \in R$ and any distribution $\dot{p} \in R^I$, define the set $X(r; \dot{p})$ of states by $X(r; \dot{p}) = \{(w, p) \in X : p = \dot{p}$ and $\sum_{i \in \dot{p}^r} w_i = 1\}$. We first show that $D \cap$

$X(r;p)$ is the unique set that satisfies both internal stability and external stability with respect to $X(r;p)$. By Theorem 25, the unique stable set in a one-region model is the set of dyadic states. Given a region $r \in R$ and a distribution $p \in R^I$, define the function $w^{r,p} : X \longrightarrow [0,1]^{\#p^r}$ by $w^{r,p}(w,p)_1 = w_{\min p^r}, ..., w^{r,p}(w,p)_{\#p^r} = w_{\max p^r}$; that is, $w^{r,p}$ projects from X onto allocations of players in the region r of the distribution p. Then $\{w^{r,p}(w,p) : (w,p) \in D \cap X(r;p)\}$ is the set of allocations of dyadic states in the $\#p^r$-player one-region model, and thus it is the unique stable set by Theorem 25 in this one-region model. Note that in a one-region model, dominance relation between states is well defined without distributions. Thus it is easily seen that $(w',p) \in X(r;p)$ dominates $(w,p) \in X(r;p)$ if and only if $w^{r,p}(w',p)$ dominates $w^{r,p}(w,p)$ in the $\#p^r$-player one-region model; both mean that $\sum_{z \in \{i : w'_i > w_i\}} w_z > \sum_{z \in \{i : w'_i < w_i\}} w_z$. Therefore, $D \cap X(r;p)$ is the unique set that satisfies both internal stability and external stability with respect to $X(r;p)$ because $\{w^{r,p}(w,p) : (w,p) \in D \cap X(r;p)\}$ is the unique stable set of allocations in the $\#p^r$-player one-region model.

For any region $r \in R$, any distribution $\dot{p} \in R^I$, and any player z with $\dot{p}_z \notin \dot{p}^r$, define the set $X(z,r;\dot{p})$ of states by $X(z,r;\dot{p}) = \{(w,p) \in X : p = \dot{p}, \sum_{i \in \dot{p}^r} w_i = \frac{1}{2}$, and $w_z = \frac{1}{2}\}$. We second prove that $D \cap X(z,r;p)$ is the unique set that satisfies both internal stability and external stability with respect to $X(z,r;p)$. Note that $(w',p) \in X(z,r;p)$ dominates $(w,p) \in X(z,r;p)$ if and only if $2w^{r,p}(w',p)$

dominates $2w^{r,p}(w,p)$ in the $\#p^r$-player one-region model. It is easily seen that $\{2w^{r,p}(w,p) : (w,p) \in D \cap X(z,r;p)\}$ is the set of allocations of dyadic states in the $\#p^r$-player one-region model, and thus by Theorem 25, $\{2w^{r,p}(w,p) : (w,p) \in D \cap X(z,r;p)\}$ is the unique stable set. Therefore, $D \cap X(z,r;p)$ is the unique set that satisfies both internal stability and external stability with respect to $X(z,r;p)$.

Third, we check that a state in $X(r;p)$ can be dominated only by another state in $X(r;p)$. If $(w,p) \in X(r;p)$ is dominated by another state (w',p'), then because the coalition $\{i : w'_i > w_i\} \subset p^r$ pillages the coalition $\{i : w'_i < w_i\} \subset p^r$ within region r, we have that $p_i = p'_i = r$ for any $i \in \{i : w'_i \neq w_i\}$. Since the pillage does not affect the coalition $\{i : w'_i = w_i\}$, we have that $p_i = p'_i$ for any $i \in \{i : w'_i = w_i\}$. Since $p' = p$ and $p'_i = r$ for each $i \in \{i : w'_i > 0\}$, we have that $(w',p') \in X(r;p)$.

Suppose that $\bar{S} \subset \bar{X}$ is a set that satisfies both internal stability and external stability with respect to $\bar{X}$. We next demonstrate that $\bar{S} = D$. The set $\bar{S}$ must dominate every state in $X(r,p) \setminus \bar{S}$. However, $X(r,p) \setminus \bar{S}$ can be dominated only by some state in $X(r,p)$, and thus $\bar{S} \cap X(r,p)$ dominates every state in $X(r,p) \setminus \bar{S}$. Since $\bar{S} \cap X(r,p)$ is internally stable, $\bar{S} \cap X(r,p)$ is a set that satisfies both internal stability and external stability with respect to $X(r,p)$. Therefore, we have that $\bar{S} \cap X(r,p) = D \cap X(r;p)$. Since r and p are arbitrary, we have that $\bar{S} \cap \bigcup_{(r,p) \in R \times R^I} X(r,p) = D \cap \bigcup_{(r,p) \in R \times R^I} X(r;p)$. Note that a state in

$X(z,r;p)$ can be dominated only by some state in $X(z,r;p) \cup \bigcup_{(r,p)\in R\times R^I} X(r;p)$. Any state in $D \cap \bigcup_{(r,p)\in R\times R^I} X(r;p)$ cannot dominate another state in $X(z,r;p)$ because a state (w',p') that results from player z with $w_z = \frac{1}{2}$ and $p_z \neq r$ pillaging other players at region r at (w,p) in $X(z,r;p)$ has $1 > w'_z > \frac{1}{2}$, and thus $(w',p') \notin D$. Therefore, $\bar{S} \cap X(z,r;p)$ must dominate every state in $X(z,r;p) \setminus \bar{S}$ because $\bar{S}$ dominates every state in $X(z,r;p) \setminus \bar{S}$. Since $\bar{S} \cap X(z,r;p)$ is internally stable, $\bar{S} \cap X(z,r;p)$ satisfies both internal stability and external stability with respect to $X(z,r;p)$. Thus we have that $\bar{S} \cap X(z,r;p) = D \cap X(z,r;p)$ because $D \cap X(z,r;p)$ is the unique set that satisfies both internal stability and external stability with respect to $X(z,r;p)$. Since r, p, and z with $\dot{p}_z \notin \dot{p}^r$ are arbitrary, we have that $\bar{S} \cap \bigcup_{(r,p)\in R\times R^I} (\bigcup_{z\notin p^r} X(z,r;p)) = D \cap \bigcup_{(r,p)\in R\times R^I} (\bigcup_{z\notin p^r} X(z,r;p))$. Since $\bigcup_{(r,p)\in R\times R^I} (X(r;p) \cup \bigcup_{z\notin p^r} X(z,r;p)) = \bar{X}$ and $\bar{S}, D \subset \bar{X}$, we have that $\bar{S} = D$.

Finally, we complete the proof that D is the unique set that satisfies both internal stability and external stability with respect to $\bar{X}$. We have proven that if a set satisfies both internal stability and external stability with respect to $\bar{X}$, then it must be D. Therefore, we need to show that D satisfies both internal stability and external stability with respect to $\bar{X}$. Because for any states $(w,p), (w',p') \in D$, we have that $\sum_{z\in\{i:w'_i>w_i\}} w_z \leq \sum_{z\in\{i:w'_i<w_i\}} w_z$ or $\sum_{z\in\{i:w_i>w'_i\}} w'_z \leq \sum_{z\in\{i:w_i<w'_i\}} w'_z$, the set D is internally stable. Note that for each r, p, and z with $p_z \notin p^r$, $D \cap X(r;p)$

and $D \cap X(z, r; p)$ satisfy external stability with respect to $X(r; p)$ and $X(z, r; p)$, respectively. Therefore, D is externally stable with respect to $\bar{X}$. Consequently, D satisfies both internal stability and external stability with respect to $\bar{X}$.

In addition, It is easily seen that a stable set S includes D. Note that a state in $\bar{X}$ can be dominated only by another state in $\bar{X}$. Every state (w', p') that results from player z with $w_z = \frac{1}{2}$ being involved in pillaging other players at (w, p) in $\bar{X}$ satisfies that $\sum_{i \in p'^r} w'_i = 1$ for some $r \in R$ and thus that $(w', p') \in \bar{X}$. Every state (w'', p'') that results from players in some region r pillaging other players in the same region r at (w, p) in $\bar{X}$ satisfies that for each $r \in R$, $\sum_{i \in p''^r} w''_i = 0, \frac{1}{2}$, or 1 and that if $\sum_{i \in p''^r} w''_i = \frac{1}{2}$ for some region $r \in R$, then $w''_z = \frac{1}{2}$ for some player z. Thus we have that $(w'', p'') \in \bar{X}$. Since a stable set S dominates every state in $\bar{X} \setminus S$, $S \cap \bar{X}$ dominates every state in $\bar{X} \setminus S$. Since $S \cap \bar{X}$ is internally stable, $S \cap \bar{X}$ satisfies both internal stability and external stability with respect to $\bar{X}$. Therefore, we have that $S \cap \bar{X} = D$ and thus that $D \subset S$. Since S is an arbitrary stable set, a stable set includes D. $\square$

Proposition 27 completely characterizes stable sets in the three-player and two-region model.

Proposition 27. *In the three-player and two-region model, a set S is a stable set if and only if $S = B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$ for some basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$.*

Proof. We prove the necessary condition first. Suppose that S is a stable set. By Lemma 23, there exist basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$ such that $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \subset S$. By Lemma 26, we have that $D \subset S$. Therefore, we must have that $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D \subset S$. To show that $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D \supset S$, it suffices to show that $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$ is externally stable.

Let $\dot{S} = B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$. If $(\mathring{w}, p) \in X \setminus \dot{S}$ with $\mathring{w}_z < \frac{1}{2}$ for each $z \in I$, then $(\mathring{w}, p)$ is dominated by some state in D such that two players have halves. Let $\bar{X} = \{(w,p) \in X :$ for each region $r \in R$, $\sum_{i \in p^r} w_i = 0, \frac{1}{2}$, or 1 and if for some region $r' \in R$, $\sum_{i \in p^{r'}} w_i = \frac{1}{2}$, then for some player z, $w_z = \frac{1}{2}\}$. If $(\mathring{w}, p) \in X \setminus \dot{S}$ and $(\mathring{w}, p) \in \bar{X}$, then by Lemma 26, $(\mathring{w}, p)$ is dominated by some state in D such that either two players have halves, or one player has all of the wealth. If $(\mathring{w}, p) \in X \setminus \dot{S}$ with $\mathring{w}_i > \frac{1}{2}$ and $p_j = p_k$, then $(\mathring{w}, p)$ is dominated by some state in D such that player i has all of the wealth.

Let $(\mathring{w}, p) \in X \setminus \dot{S}$ satisfy that $p \in B(i; j, k)$ and $(\mathring{w}, p) \in B^i_{j,k}(\frac{1}{2}, p)$. Without loss of generality, we assume that $p_i = p_j$. Since $1 > \mathring{w}_i \geq \frac{1}{2}$ and a basic function

$\beta^i_{j,k}(\cdot,p) : [0,\frac{1}{2}] \rightarrow [0,\frac{1}{2}]$ is uniformly continuous on $[0,\frac{1}{2})$, $\lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$ is well defined. If $\mathring{w}_i > \frac{1}{2}$ and $\mathring{w}_j > \beta^i_{j,k}(1-\mathring{w}_i,p)$ or $\mathring{w}_i = \frac{1}{2}$ and $\mathring{w}_j > \lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$, both of which mean that $\mathring{w}_j > \lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$, then $\mathring{w}_k < 1-\mathring{w}_i-\lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$. Thus there exists a state $(w,p) \in B(\beta^i_{j,k})$ with $w_i > \mathring{w}_i$ and $w_k = \mathring{w}_k$. In this case, (w,p) dominates $(\mathring{w},p)$ by player i pillaging player j. If $\mathring{w}_i > \frac{1}{2}$ and $\mathring{w}_j < \beta^i_{j,k}(1-\mathring{w}_i,p)$ or $\mathring{w}_i = \frac{1}{2}$ and $\mathring{w}_j < \lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$, both of which mean that $\mathring{w}_j < \lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$, then $(w',\dot{p}) \in B(\beta^i_{j,k})$ such that $w'_j = \lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p)$, $\dot{p}_z = p_z$ for all $z \neq i$, and $\dot{p}_i \neq p_i$ dominates $(\mathring{w},p)$ by player i pillaging player k. If $\mathring{w}_i = \frac{1}{2}$ and $\mathring{w}_j = \lim_{\lambda\longrightarrow 1-\mathring{w}_i} \beta^i_{j,k}(\lambda,p) \neq \frac{1}{4}$, then some state in the core such that two players have halves dominates $(\mathring{w},p)$. Therefore, $\dot{S}$ is externally stable, and thus $\dot{S} = S$.

Next, we prove the sufficient condition, that is, if functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$ are basic functions, then the set $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$ is a stable set. Suppose that $S' = B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$ for some basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$. Then S' is externally stable as shown above. Now, we need to show that S' is internally stable.

Notice that each $B(i;j,k)$ has four elements and each element p in $B(i;j,k)$ has its counterpart distribution $\dot{p}$ such that $\dot{p} \in B(i;j,k)$, $p_i \neq \dot{p}_i$, $\dot{p}_j = p_j$, and $\dot{p}_k = p_k$. For example, $B(1;2,3) = \{(1,1,2),(1,2,1),(2,1,2),(2,2,1)\}$ and $(1,1,2)$

and $(1,2,1)$ are counterpart distributions to $(2,1,2)$ and $(2,2,1)$, respectively. Therefore, by the second condition of a basic function, if $(w,p),(w',p) \in B(\beta^i_{j,k})$ with $w_i > w'_i > \frac{1}{2}$ and $p_i = p_j$, then $(w,\dot{p}),(w',\dot{p}) \in B(\beta^i_{j,k})$ with $\dot{p}_i = \dot{p}_k$, and vice versa. Then we have that $w_j \leq w'_j$ and $w_k < w'_k$ by the first condition of a basic function since $p_i = p_j$ and $\dot{p}_i = \dot{p}_k$. Similarly, we have that $w_j < w'_j$ and $w_k \leq w'_k$. Consequently, we have that $w_j < w'_j$ and $w_k < w'_k$.

First, we prove that each $B(\beta^1_{2,3})$, $B(\beta^2_{3,1})$, and $B(\beta^3_{1,2})$ is internally stable. Let $(w,p),(w',p') \in B(\beta^i_{j,k})$ such that $(w,p) \neq (w',p')$ and $(w,p) \notin C$, which is the core. Since $\{p_j,p_k\} = \{p'_j,p'_k\} = R$, i.e., players j and k are distributed all over regions at p and p', we have that either $p'_j \neq p_j$ and $p'_k \neq p_k$, or $p'_j = p_j$ and $p'_k = p_k$. Thus if $p'_j \neq p_j$, then $\{z : p_z \neq p'_z\} \nsubseteq p^r$ for each $r \in R$, and so (w',p') does not dominate (w,p).

Suppose that p and p' satisfies that $p'_j = p_j$ and $p'_k = p_k$. If $w_i, w'_i > \frac{1}{2}$, then *i)* $w_j < w'_j$ and $w_k < w'_k$; *ii)* $w_j > w'_j$ and $w_k > w'_k$; or *iii)* $w = w'$. If either *i)* $w_j < w'_j$ and $w_k < w'_k$, or *ii)* $w_j > w'_j$ and $w_k > w'_k$, then $\{z : w'_z \neq w_z\} \nsubseteq p^r$ for each $r \in R$. If $w = w'$ then $\sum_{z \in \{i : w'_i > w_i\}} w_z = \sum_{z \in \{i : w'_i < w_i\}} w_z = 0$. If $w_i > \frac{1}{2}$ and $w'_i = \frac{1}{2}$, then $\sum_{z \in \{i : w'_i > w_i\}} w_z \leq \frac{1}{2} < \sum_{z \in \{i : w'_i < w_i\}} w_z$. If $w_i = \frac{1}{2}$, $w_j = \frac{1}{4}$, and $w'_i > \frac{1}{2}$, then since $\lim_{\lambda \longrightarrow \frac{1}{2}} \beta^i_{j,k}(\lambda) = \frac{1}{4}$, $w_j = \frac{1}{4} > w'_j$ and $w_k = \frac{1}{4} > w'_k$. Thus we have that $\{i : w'_i \neq w_i\} \nsubseteq p^r$ for each $r \in R$. If $w_i = \frac{1}{2}$, $w_j = \frac{1}{4}$, and $w'_i = \frac{1}{2}$, then since

$w'_j = \frac{1}{4}$ or $\frac{1}{2}$, we have that $\sum_{z\in\{i:w'_i>w_i\}} w_z = \sum_{z\in\{i:w'_i<w_i\}} w_z = 0$ or $\frac{1}{4}$. Therefore, in these cases, (w', p') does not dominate (w, p). Since $(w, p), (w', p') \in B(\beta^i_{j,k})$ with $(w, p) \neq (w', p')$ and $(w, p) \notin C$ are arbitrary, each set $B(\beta^1_{2,3})$, $B(\beta^2_{3,1})$, and $B(\beta^3_{1,2})$ is internally stable.

Second, we check internal stability of the set $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2})$. Let $(w, p) \in B(\beta^i_{j,k})$ and $(w'', p'') \in B(\beta^j_{k,i})$. Then $w_i \geq \frac{1}{2}$ and $w''_j \geq \frac{1}{2}$, and thus $w_j \leq \frac{1}{2}$ and $w''_i \leq \frac{1}{2}$. If $w_i > w''_i$ then $\sum_{z\in\{y:w''_y>w_y\}} w_z \leq \frac{1}{2} \leq w_i \leq \sum_{z\in\{y:w''_y<w_y\}} w_z$. If $w_i = w''_i$ then $w_i = w''_i = \frac{1}{2}$. Since $w_j \in \{\frac{1}{4}, \frac{1}{2}\}$ by the third condition of a basic function, $\sum_{z\in\{i:w''_i>w_i\}} w_z = \sum_{z\in\{i:w''_i<w_i\}} w_z = 0$ or $\frac{1}{4}$. Therefore, (w, p) does not dominate (w'', p''). Similarly, we can prove that (w'', p'') does not dominate (w, p). Consequently, $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2})$ is internally stable.

Finally, we examine internal stability of $S' = B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$. Note that $(B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2})) \cap \bar{X} \subset C$ and that $C \subset D$. Since a state in D can be dominated only by another state in $\bar{X}$ and D is internally stable, any state in D is not dominated by another state in $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2})$. Therefore, it suffices to show that any state in $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2})$ is not dominated by another state in D. Let $(w, p) \in B(\beta^i_{j,k})$. Then we have that $w_i \geq \frac{1}{2}$. If $w_i = 1$ or $w_i = \frac{1}{2}$ and $w_j = \frac{1}{2}$, then $(w, p) \in C$, and thus (w, p) is not dominated by any state in D. If $1 > w_i > \frac{1}{2}$, then by the basic condition of $B(\beta^i_{j,k})$, we have that $w_j > 0$ and $w_k > 0$. Also, if $w_i = \frac{1}{2}$ and $w_j = \frac{1}{4}$, then $w_k = \frac{1}{4}$. In these cases, player i cannot pillage both players j and k simultaneously since $p_j \neq p_k$ and cannot be pillaged by another player since $w_i > \frac{1}{2} > \max\{w_j, w_k\}$. Thus if a state (w''', p''') dominates (w, p), then (w''', p''') satisfies that for some player z, $w_z \notin \{0, \frac{1}{4}, \frac{1}{2}, 1\}$, that is, $(w''', p''') \notin D$. Since (w, p) is arbitrary, any state in $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2})$ is not dominated by another state in D.

Therefore, S' is a stable set. Since functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$ are arbitrary basic functions, $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$ for any basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$ is a stable set. $\square$

Figures 2.3 and 2.4 show one possible stable set S on the hyperplanes. Dots and bold curves in the figures denote states in S at each distribution. As shown in the proof of Proposition 27, the bold curves in the figures can be expressed with basic functions.

Theorem 28 generalizes Proposition 27 to the three-player $N-$region models where $N \geq 2$.

Theorem 28. *In a three-player $N-$region model where $N \geq 2$, a set S is a stable set if and only if $S = B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D \cup U(B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup C)$ for some basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$ where the set*

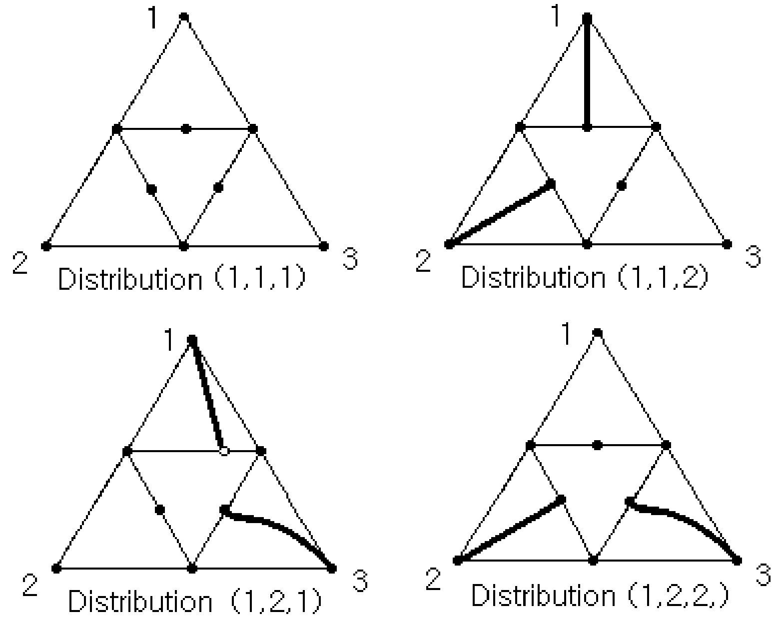

Figure 2.3. Stable set in the hyperplanes of states with $p_1 = 1$

$U(B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup C)$ is the set of states that are not dominated by any state in $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup C$.

Proof. For any two distinct regions $q, r \in R$, define the set $X^{q,r}$ of states by $X^{q,r} = \{(w,p) \in X : \text{for each } i,\ p_i = q \text{ or } r\}$. Then, it is easily seen that a state in $X^{r,q}$ can be dominated only by some state in $X^{r,q}$ because the act of the pillage does not disperse players. If there are more than two regions, then define the set $X^{indiv.}$ of states by $X^{indiv.} = \{(w,p) \in X : \text{for any three distinct regions } o, q, \text{ and } r,\ \{p_1,p_2,p_3\} = \{o,q,r\}\}$. That is, $X^{indiv.}$ is the set of states at which each player occupies its own region alone, i.e., individual region distribution. Note

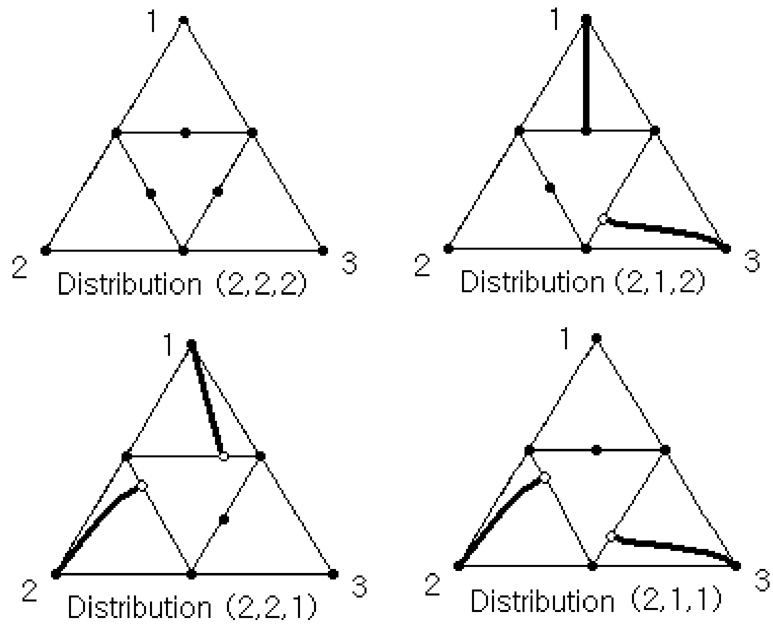

Figure 2.4. Stable set in the hyperplanes of states with $p_1 = 2$

that any state in $X^{indiv.}$ does not dominate any other state in X, however, it can be dominated by some state at which only one region contains two players, whose distribution results from one player pillaging another player. Therefore, a set S is a stable set if and only if i) for any two distinct regions q and r, $S \cap X^{r,q}$ is both internally stable and externally stable with respect to $X^{r,q}$; and ii) S dominates all states in $X^{indiv.}$ except itself $X^{indiv.} \cap S$.

In the three-player and two-region model, by Proposition 27, a set S is a stable set if and only if $S = B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D$ for some basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$. Without loss of generality, given any two distinct regions q and r, we can regard a state (w,p) in $X^{r,q}$ as the state (w,p) in the two-region model, and vice versa. Then, it is easily seen that $(w',p') \in X^{r,q}$ dominates $(w,p) \in X^{r,q}$ if and only if (w',p') dominates (w,p) in the two-region model. Therefore, for any two distinct regions q and r, $S \cap X^{r,q}$ is both internally stable and externally stable with respect to $X^{r,q}$ if and only if $S \cap X^{r,q} = (B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup D) \cap X^{r,q}$ for some basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$. The observation that for any basic functions $\beta^1_{2,3}$, $\beta^2_{3,1}$, and $\beta^3_{1,2}$, the set $B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup C$ dominates all states in $X^{indiv.}$ except $U(B(\beta^1_{2,3}) \cup B(\beta^2_{3,1}) \cup B(\beta^3_{1,2}) \cup C)$ completes the proof. □

Figures 2.5 and 2.6 show one possible stable set S on the hyperplanes. Figure 2.5 covers distributions where at least two players are in a common region and Figure 2.6 covers the other distributions, where each player occupies its own region alone. In the figures, dots, bold lines, and the gray area denote states in the stable set S. Note that except for the three corner points and three middle points, the gray area does not contain boundary lines.

2.2.2. Stable set in $I-$player and $N-$region models where $I = 4$ and $N = 2$, or $I \geq 4$ and $N \geq 3$

A stable set does not exist in a $I-$player and $N-$region model where $I = 4$ and $N = 2$, or $I \geq 4$ and $N \geq 3$. First, we prove that in the four-player and two-region

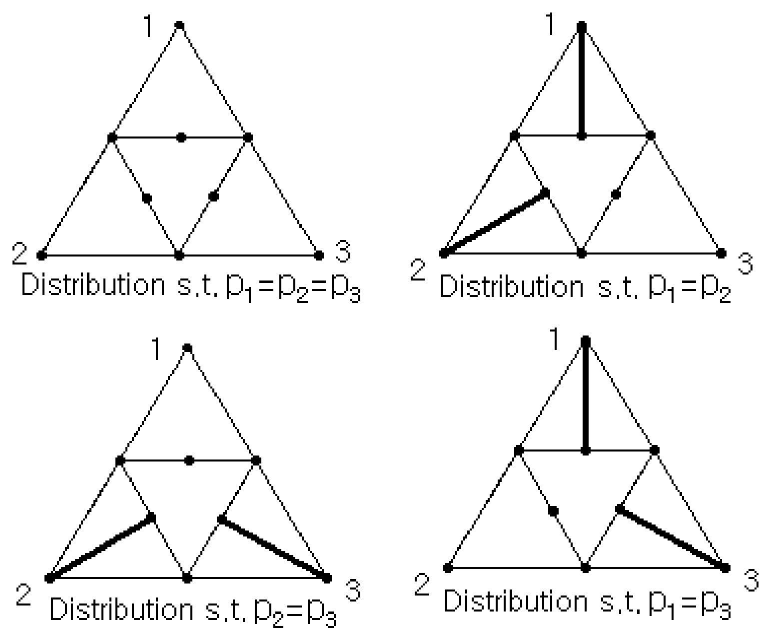

Figure 2.5. Stable set in the hyperplanes of states such that $p_j = p_k$ for two distinct players j and k.

model, a stable set must contain a group of states out of basic sets. Second, we discover some properties of a group of states that are not in basic sets, but are in a stable set. Next, we show that in the four-player and two-region model, if there exists a stable set S, then we can find a state (w, p) such that (w, p) cannot be in S and S cannot dominate (w, p). It is because S contains four states and the properties of these states assure that S dominates every state that dominates (w, p). Finally, we generalize the result in the four-player and two-region model and verify the nonexistence of a stable set in a $I-$player and $N-$region model where $I \geq 4$ and $N \geq 3$.

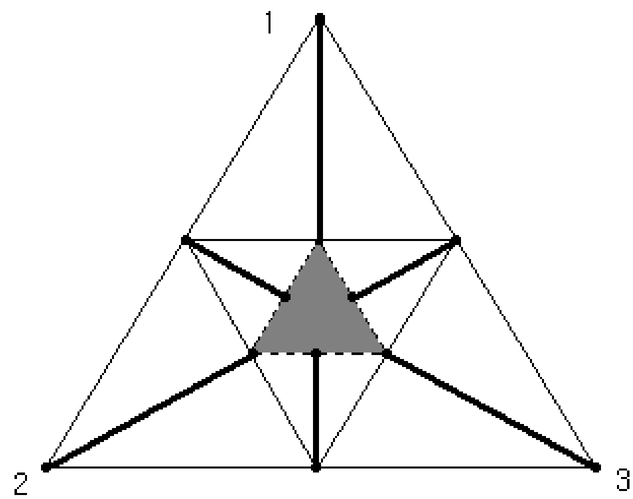

Figure 2.6. Stable set in the hyperplane of states such that $p_1 \neq p_2 \neq p_3 \neq p_1$

Lemma 29 shows another property of a stable set. To satisfy both internal and external stabilities, a stable set must contain some states outside the basic sets as well as some basic states. Lemma 29 reveals relation among states that are outside the basic states and belong to some stable set.

Lemma 29. *In the four-player and two-region model, for any player j, let a distribution p satisfy $p^1 = \{j\}$ or $p^2 = \{j\}$. Then given a stable set S, there exists a positive real number a_p such that $[0, a_p] \subset \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$. In particular, if $(w', p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$, then $[0, w'_j] \subset \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$.*

Proof. Let a distribution $\dot{p} \in B(i; j, k)$ with $\dot{p}_z = p_z$ for each $z \neq i$. Note that $B(i; j, k)$ is the set of distributions at which player i is together with only either player j or player k; that is, at the distribution $\dot{p}$, there exists a region $r \in R$ such that $\dot{p}^r = \{i, j\}$ or $\{i, k\}$. At the distribution p, player j is alone in a region and player k is together with the other players including player i. Therefore, we must have that $\dot{p}_i = \dot{p}_j$ so that player i is together with only one player in a common region. By Lemma 23, there exists a basic function $\beta^i_{j,k}$ with $(B(\beta^i_{j,k}) \cup C) \cap B^i_{j,k}(\frac{1}{2}, \dot{p}) = S \cap B^i_{j,k}(\frac{1}{2}, \dot{p})$. By the first condition of a basic function, there exists $\dot{\lambda} \in (0, \frac{1}{2})$ with $\beta^i_{j,k}(\dot{\lambda}, \dot{p}) > 0$. According to Corollary 22, $\beta^i_{j,k}(\cdot, \dot{p})$ is uniformly continuous on $[0, \frac{1}{2})$. Since $\beta^i_{j,k}(0, \dot{p}) = 0$, the intermediate value theorem implies that $[0, \beta^i_{j,k}(\dot{\lambda}, \dot{p})] \subset \beta^i_{j,k}([0, \dot{\lambda}], \dot{p})$.

To prove the first assertion, it suffices to show that $[0, \beta^i_{j,k}(\dot{\lambda}, \dot{p})] \subset \{w_j : (w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$. For any $\delta \in [0, \beta^i_{j,k}(\dot{\lambda}, \dot{p})]$, let $\lambda_\delta = \min\{\lambda \in [0, \dot{\lambda}] : \beta^i_{j,k}(\lambda, \dot{p}) = \delta\}$, which is well defined by the uniform continuity of $\beta^i_{j,k}$ on $[0, \dot{\lambda}]$ because if a function is continuous, then the inverse image of a closed set under the function is a closed set. Given any $\lambda_\delta \in [0, \dot{\lambda}]$, let the allocation w^δ satisfy that $w^\delta_i = 1 - \lambda_\delta$, $w^\delta_j = \beta^i_{j,k}(\lambda_\delta, \dot{p})$, and $w^\delta_k = \lambda_\delta - \beta^i_{j,k}(\lambda_\delta, \dot{p})$. Then we have that $(w^\delta, \dot{p}) \in B(\beta^i_{j,k}) \cap B^i_{j,k}(\frac{1}{2}, \dot{p})$ and thus that $(w^\delta, \dot{p}) \in S$. Suppose by way of contradiction that $(w^\delta, p) \notin S$. Every state that results from player i pillaging either player j or player k at (w^δ, p) dominates $(w^\delta, \dot{p}) \in S$ as well as (w^δ, p). Every state that results from either player j or player k engaging in pillage at (w^δ, p), regardless of player $i's$ participation, is dominated by some state in the core such that player i has the total wealth because players j and k get together in a common region and player i has greater than a half. Therefore, S cannot dominate (w^δ, p), and thus S lacks external stability. This contradiction guarantees that $(w^\delta, p) \in S$ and thus that $\delta = w_j^\delta \in \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$ since $(w^\delta, p) \in B^i_{j,k}(\frac{1}{2}, p)$. Since $\delta \in [0, \beta^i_{j,k}(\dot{\lambda}, \dot{p})]$ is arbitrary, we have that $[0, \beta^i_{j,k}(\dot{\lambda}, \dot{p})] \subset \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$.

In particular, if $(w', p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$, then by Lemma 18, $(w', \dot{p}) \in S \cap B^i_{j,k}(\frac{1}{2}, \dot{p})$. By the same way as shown above, we can show that for any $\delta \in [0, w'_j]$, $\delta \in \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$ and thus that $[0, w'_j] \subset \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$. □

Lemmas 30, 31, and 32 strengthen Lemma 29 by revealing relations between states that are outside the basic states and belong to some stable set.

Lemma 30. *In the four-player and two-region model, for any player j, let a distribution p satisfy either $p^1 = \{j\}$ or $p^2 = \{j\}$. Then given a stable set S, for any $a \in (0, \frac{1}{2})$, there exists a state $(w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ such that $a > w_j + w_k > w_j > 0$ and $[0, w_j] \subset \{w_j : (w,p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$. In addition, if $(w', p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ with $w_j > w'_j > 0$, then we have that $w_k \geq w'_k > 0$.*

Proof. By Lemma 23, there exists a basic function $\beta^i_{j,k}$ with $B(\beta^i_{j,k}) \subset S$. Let the distribution $\dot{p} \in B(i; j, k)$ satisfy $\dot{p}_z = p_z$ for each $z \neq i$. Then since $\beta^i_{j,k}(\cdot, \dot{p})$

is defined on $[0, \frac{1}{2}]$, for any $a \in (0, \frac{1}{2})$, we can find the allocation w^a with $w^a_j = \beta^i_{j,k}(a, \dot{p})$, $w^a_k = a - \beta^i_{j,k}(a, \dot{p})$, and $w^a_i = 1 - a$. Then we have that $(w^a, \dot{p}) \in B(\beta^i_{j,k})$. By Lemma 29, there exists a state $(\dot{w}, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ with $\dot{w}_j > 0$ and $[0, \dot{w}_j] \subset \{w_j : (w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$. Let a state $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ satisfy that $0 < w_j < \min\{\dot{w}_j, w^a_j\}$. We will show that (w, p) satisfies all required conditions, that is, $a > w_j + w_k > w_j > 0$ and $[0, w_j] \subset \{w_j : (w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$.

Since $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$, by Lemma 18, we have that $(w, \dot{p}) \in S \cap B^i_{j,k}(\frac{1}{2}, \dot{p})$. Since $w^a_j > w_j > 0$ and $w^a_i > \frac{1}{2}$, the basic condition means that $w_i > w^a_i$ and $w_k > 0$, and thus we have that $a > 1 - w_i = w_j + w_k > w_j > 0$. Since $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$, Lemma 29 assures the second condition, $[0, w_j] \subset \{w_j : (w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)\}$.

In addition, let $(w', p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ with $w'_j \in (0, w_j)$. Then by Lemma 18, we have that $(w', \dot{p}) \in S \cap B^i_{j,k}(\frac{1}{2}, \dot{p})$. Since $w_j > w'_j > 0$ and $w_i > 1 - a > \frac{1}{2}$, the basic condition of S implies that $w_i < w'_i < 1$ and thus that $w_k \geq w'_k > 0$. $\square$

Lemma 31. *In the four-player and two-region model, for the distinct four players i, j, k, and y, let distributions p and p' satisfy either $p^1 = \{j\}$ and $p'^1 = \{i, j, y\}$, or $p^2 = \{j\}$ and $p'^2 = \{i, j, y\}$. Given a stable set S, suppose that $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ and $(w', p') \in S \cap B^i_{k,y}(\frac{1}{2}, p')$. If $w_i, w'_i > \frac{1}{2}$ and $w_j = w'_k$, then $w_k \geq w_j$ or $w'_y \geq w'_k$.*

Proof. By way of contradiction, suppose that $w_i, w'_i > \frac{1}{2}$, $w_j = w'_k$, $w_j > w_k$, and $w'_k > w'_y$. Lemma 29 implies that $[0, w'_k] \subset \{w_k : (w, p') \in S \cap B^i_{k,y}(\frac{1}{2}, p')\}$. Since $w'_k = w_j > w_k$, there exists a state $(\dot{w}', p') \in S \cap B^i_{k,y}(\frac{1}{2}, p')$ with $\dot{w}'_k = w_k$. Let $\dot{p}' \in B(i; k, y)$ satisfy $\dot{p}'_i \neq p'_i$ and $\dot{p}'_z = p'_z$ for each $z \neq i$. Since $(w', p'), (\dot{w}', p') \in S \cap B^i_{k,y}(\frac{1}{2}, p')$, we have that $(w', \dot{p}'), (\dot{w}', \dot{p}') \in S \cap B^i_{k,y}(\frac{1}{2}, \dot{p}')$ by Lemma 18. Since $w_i, w'_i > \frac{1}{2}$ and $w'_k > w_k = \dot{w}'_k$, the basic condition of S means that $w'_i < \dot{w}'_i$ and thus that $w'_y \geq \dot{w}'_y$. Since $w'_k > w'_y$, we have that $w_j = w'_k > w'_y \geq \dot{w}'_y$ and thus that $\dot{w}'_i = 1 - \dot{w}'_k - \dot{w}'_y > 1 - \dot{w}'_k - w_j = 1 - w_k - w_j = w_i$. Since $w_i + w_j + w_k = 1$, we have that $w_y = 0$. Therefore, we have that $\dot{w}'_i > w_i > \frac{1}{2}$, $\dot{w}'_k = w_k$, and $\dot{w}'_y > w_y = 0$. Note that the distribution p' results from players i and y moving to the region of player j at the distribution p. Therefore, $(\dot{w}', p') \in S$ dominates $(w, p) \in S$ by players i and y pillaging player j at (w, p). This contradiction guarantees that if $w_i, w'_i > \frac{1}{2}$ and $w_j = w'_k$, then $w_k \geq w_j$ or $w'_y \geq w'_k$. □

Lemma 32. *In the four-player and two-region model, for any player j, let a distribution p satisfy either $p^1 = \{j\}$ or $p^2 = \{j\}$. Given a stable set S, if $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$ and $(w', p) \in S \cap B^i_{j,y}(\frac{1}{2}, p)$ with $w_j = w'_j > 0$, then $w_i = w'_i$.*

Proof. Suppose by way of contradiction that $(w, p) \in S \cap B^i_{j,k}(\frac{1}{2}, p)$, $(w', p) \in S \cap B^i_{j,y}(\frac{1}{2}, p)$, $w_j = w'_j > 0$ and $w_i \neq w'_i$. Since $(w, p) \in B^i_{j,k}(\frac{1}{2}, p)$ and $(w', p) \in B^i_{j,y}(\frac{1}{2}, p)$, we have that $w_i \geq \frac{1}{2} \geq w_j + w_k$ and $w'_i \geq \frac{1}{2} \geq w'_j + w'_y$. Since $w_j, w'_j > 0$, we have that $w_i > w_k$ and $w'_i > w'_y$. Therefore, if $w_i > w'_i$ then $(w,p) \in S$ dominates $(w',p) \in S$ by either player i or players i and k pillaging player y at (w',p). Similarly, if $w_i < w'_i$ then $(w',p) \in S$ dominates $(w,p) \in S$. This contradiction completes the proof. $\square$

Lemma 33 synthesizes the previous results in this subsection and shows that in the four-player and two-region model, a stable set S contains four distinct states that satisfy six conditions introduced in this lemma. We can use these four states to show nonexistence of stable set. The six conditions guarantee that there exists a state (w,p) that S cannot contain or dominate.

Lemma 33. *In the four-player and two-region model, a stable set S contains four states $(\dot{w},p')$, $(\ddot{w},p'')$, $(\dddot{w},p''')$, and $(\mathring{w},p''')$ such that for some four distinct players i, j, k, and y, i) distributions p', p'', and p''' satisfy either $p'^1 = \{j\}$, $p''^1 = \{i,j,y\}$, and $p'''^1 = \{k\}$, or $p'^2 = \{j\}$, $p''^2 = \{i,j,y\}$, and $p'''^2 = \{k\}$; ii) $(\dot{w},p') \in S \cap B^i_{j,k}(\frac{1}{2},p')$, $(\ddot{w},p'') \in S \cap B^i_{k,j}(\frac{1}{2},p'')$, and $(\dddot{w},p'''),(\mathring{w},p''') \in S \cap B^i_{k,j}(\frac{1}{2},p''')$; iii) $\ddot{w}_k > \dot{w}_k \geq \dot{w}_j > 0$; iv) $\ddot{w}_j, \dddot{w}_j < \frac{1}{4}$; v) $\frac{1}{4} > \dddot{w}_k > \dot{w}_k + \dot{w}_j$; and vi) $\mathring{w}_j \geq \mathring{w}_k = \dot{w}_j$.*

Proof. Let distributions p, $p(1)$, $p(2)$, $p(3)$, $p(4)$, and $p(5)$ satisfy that either $p^1 = \{i,j,y\}$, $p(1)^1 = \{j\}$, $p(2)^1 = \{i,j,k\}$, $p(3)^1 = \{k\}$, $p(4)^1 = \{i,k,y\}$, and $p(5)^1 = \{y\}$; or $p^2 = \{i,j,y\}$, $p(1)^2 = \{j\}$, $p(2)^2 = \{i,j,k\}$, $p(3)^2 = \{k\}$, $p(4)^2 =$

$\{i,k,y\}$, and $p(5)^2=\{y\}$. By Lemma 30, there exist states $(\ddot{w},p)\in S\cap B^i_{k,j}(\frac{1}{2},p)$, $(\ddot{w}^{(1)},p(1))\in S\cap B^i_{j,y}(\frac{1}{2},p(1))$, $(\dddot{w}^{(2)},p(2))\in S\cap B^i_{y,k}(\frac{1}{2},p(2))$, $(\dddot{w}^{(3)},p(3))\in S\cap B^i_{k,j}(\frac{1}{2},p(3))$, $(\dddot{w}^{(4)},p(4))\in S\cap B^i_{j,y}(\frac{1}{2},p(4))$, and $(\ddot{w}^{(5)},p(5))\in S\cap B^i_{y,k}(\frac{1}{2},p(5))$ with $0<\ddot{w}_j,\ddot{w}_k,\ddot{w}^{(1)}_j,\ddot{w}^{(1)}_y,\dddot{w}^{(2)}_k,\dddot{w}^{(2)}_y,\dddot{w}^{(3)}_j,\dddot{w}^{(3)}_k,\dddot{w}^{(4)}_j,\dddot{w}^{(4)}_y,\ddot{w}^{(5)}_k,\ddot{w}^{(5)}_y<\frac{1}{4}$. Lemma 30 also implies that there exist states $(\dot{w},p)\in S\cap B^i_{k,y}(\frac{1}{2},p)$, $(\mathring{w},p)\in S\cap B^i_{k,j}(\frac{1}{2},p)$, $(\dot{w}^{(1)},p(1))\in S\cap B^i_{j,k}(\frac{1}{2},p(1))$, $(\mathring{w}^{(1)},p(1))\in S\cap B^i_{j,y}(\frac{1}{2},p(1))$, $(\dot{w}^{(2)},p(2))\in S\cap B^i_{y,j}(\frac{1}{2},p(2))$, $(\mathring{w}^{(2)},p(2))\in S\cap B^i_{y,k}(\frac{1}{2},p(2))$, $(\mathring{w}^{(3)},p(3))\in S\cap B^i_{k,j}(\frac{1}{2},p(3))$, and $(\mathring{w}^{(4)},p(4))\in S\cap B^i_{j,y}(\frac{1}{2},p(4))$ such that $0<\dot{w}_k+\dot{w}_y,\dot{w}^{(1)}_j+\dot{w}^{(1)}_k,\dot{w}^{(2)}_j+\dot{w}^{(2)}_y$ $<\min\{\ddot{w}_k,\ddot{w}^{(1)}_j,\ddot{w}^{(5)}_y,\dddot{w}^{(2)}_y,\dddot{w}^{(3)}_k,\dddot{w}^{(4)}_j\}$ and $\dot{w}_k=\mathring{w}_k=\dot{w}^{(1)}_j=\mathring{w}^{(1)}_j=\dot{w}^{(2)}_y=\mathring{w}^{(2)}_y=\mathring{w}^{(3)}_k=\mathring{w}^{(4)}_j>0$.

Since $\dot{w}^{(1)}_j=\mathring{w}^{(1)}_j>0$, Lemma 32 means that $\dot{w}^{(1)}_i=\mathring{w}^{(1)}_i$ and thus that $\dot{w}^{(1)}_k=\mathring{w}^{(1)}_y$. If $\mathring{w}^{(1)}_y\geq\mathring{w}^{(1)}_j$ then $\dot{w}^{(1)}_k=\mathring{w}^{(1)}_y\geq\mathring{w}^{(1)}_j=\dot{w}^{(1)}_j$. Therefore, if $\mathring{w}^{(1)}_y\geq\mathring{w}^{(1)}_j$ and $\mathring{w}^{(3)}_j\geq\mathring{w}^{(3)}_k$, then the states $(\dot{w}^{(1)},p(1))$, $(\ddot{w},p)$, $(\dddot{w}^{(3)},p(3))$, and $(\mathring{w}^{(3)},p(3))$ satisfy all required conditions; that is, for some four distinct players i, j, k, and y, *i*) distributions $p(1)$, p, and $p(3)$ satisfy either $p(1)^1=\{j\}$, $p^1=\{i,j,y\}$, and $p(3)^1=\{k\}$, or $p(1)^2=\{j\}$, $p^2=\{i,j,y\}$, and $p(3)^2=\{k\}$; *ii*) $(\dot{w}^{(1)},p(1))\in S\cap B^i_{j,k}(\frac{1}{2},p(1))$, $(\ddot{w},p)\in S\cap B^i_{k,j}(\frac{1}{2},p)$, and $(\dddot{w}^{(3)},p(3)),(\mathring{w}^{(3)},p(3))\in S\cap B^i_{k,j}(\frac{1}{2},p(3))$; *iii*) $\ddot{w}_k>\dot{w}^{(1)}_k\geq\dot{w}^{(1)}_j>0$; *iv*) $\ddot{w}_j,\dddot{w}^{(3)}_j<\frac{1}{4}$; *v*) $\frac{1}{4}>\dddot{w}^{(3)}_k>\dot{w}^{(1)}_j+\dot{w}^{(1)}_k$; and *vi*) $\mathring{w}^{(3)}_j\geq\mathring{w}^{(3)}_k=\dot{w}^{(1)}_j$.

Note that if $\mathring{w}_i = \frac{1}{2}$, then since $\mathring{w}_j + \mathring{w}_k = \frac{1}{2}$ and $\mathring{w}_k < \frac{1}{4}$, we have that $\mathring{w}_j > \frac{1}{4} > \mathring{w}_k$, and thus some state in the core such that players i and j have halves dominates $(\mathring{w}, p) \in S$. This contradiction shows that we must have that $\mathring{w}_i > \frac{1}{2}$. Similarly, we can show that we have that $\mathring{w}_i^{(1)}$, $\mathring{w}_i^{(2)}$, $\mathring{w}_i^{(3)}$, and $\mathring{w}_i^{(4)} > \frac{1}{2}$. Therefore, if $\mathring{w}_y^{(1)} < \mathring{w}_j^{(1)}$, then since $\mathring{w}_i$, $\mathring{w}_i^{(1)}$, and $\mathring{w}_i^{(2)} > \frac{1}{2}$, Lemma 31 implies that $\mathring{w}_j \geq \mathring{w}_k$ and $\mathring{w}_k^{(2)} \geq \mathring{w}_y^{(2)}$. Since $\dot{w}_k = \mathring{w}_k$, by Lemma 32, we have that $\dot{w}_i = \mathring{w}_i$ and thus that $\dot{w}_y = \mathring{w}_j \geq \mathring{w}_k = \dot{w}_k$. In this case, the states $(\dot{w}, p)$, $(\ddot{w}^{(5)}, p(5))$, $(\dddot{w}^{(2)}, p(2))$, and $(\mathring{w}^{(2)}, p(2))$ satisfy all six conditions; that is, for some four distinct players i, k, y, and j, *i)* distributions p, $p(5)$, and $p(2)$ satisfy $p^1 = \{k\}$, $p(5)^1 = \{i, j, k\}$, and $p(2)^1 = \{y\}$, or $p^2 = \{k\}$, $p(5)^2 = \{i, j, k\}$, and $p(2)^2 = \{y\}$; *ii)* $(\dot{w}, p) \in S \cap B^i_{k,y}(\frac{1}{2}, p)$, $(\ddot{w}^{(5)}, p(5)) \in S \cap B^i_{y,k}(\frac{1}{2}, p(5))$, and $(\dddot{w}^{(2)}, p(2))$, $(\mathring{w}^{(2)}, p(2)) \in S \cap B^i_{y,k}(\frac{1}{2}, p(2))$; *iii)* $\ddot{w}_y^{(5)} > \dot{w}_y \geq \dot{w}_k > 0$, *iv)* $\ddot{w}_k^{(5)}, \dddot{w}_k^{(2)} < \frac{1}{4}$, *v)* $\frac{1}{4} > \dddot{w}_y^{(2)} > \dot{w}_k + \dot{w}_y$, and *vi)* $\mathring{w}_k^{(2)} \geq \mathring{w}_y^{(2)} = \dot{w}_k$. Similarly, if $\mathring{w}_j^{(3)} < \mathring{w}_k^{(3)}$, then $(\dot{w}^{(2)}, p(2))$, $(\ddot{w}^{(1)}, p(1))$, $(\dddot{w}^{(4)}, p(4))$, and $(\mathring{w}^{(4)}, p(4))$ satisfy all six conditions. $\square$

Proposition 34 proves nonexistence of stable set in the four-player and two-region model.

Proposition 34. *No stable set exists in the four-player and two-region model.*

Proof. By way of contradiction, suppose that there exists a stable set S in the four-player and two-region model. Then by Lemma 33, there exist four states

$(\dot{w}, p')$, $(\ddot{w}, p'')$, $(\dddot{w}, p''')$, and $(\mathring{w}, p''')$ such that for some four distinct players i, j, k, and y, $i)$ distributions p', p'', and p''' satisfy either $p'^1 = \{j\}$, $p''^1 = \{i, j, y\}$, and $p'''^1 = \{k\}$, or $p'^2 = \{j\}$, $p''^2 = \{i, j, y\}$, and $p'''^2 = \{k\}$; $ii)$ $(\dot{w}, p') \in S \cap B^i_{j,k}(\frac{1}{2}, p')$, $(\ddot{w}, p'') \in S \cap B^i_{k,j}(\frac{1}{2}, p'')$, and $(\dddot{w}, p''')$, $(\mathring{w}, p''') \in S \cap B^i_{k,j}(\frac{1}{2}, p''')$; $iii)$ $\ddot{w}_k > \dot{w}_k \geq \dot{w}_j > 0$; $iv)$ $\ddot{w}_j, \dddot{w}_j < \frac{1}{4}$; $v)$ $\frac{1}{4} > \dddot{w}_k > \dot{w}_k + \dot{w}_j$; and $vi)$ $\mathring{w}_j \geq \mathring{w}_k = \dot{w}_j$.

Define the set of states $T(\dot{w}, \ddot{w}, \dddot{w}; p') = \{(w, p) : w_i = \frac{1-\dot{w}_j}{2}, w_j = \dot{w}_j, \min\{\ddot{w}_k, \dddot{w}_k - \dot{w}_j\} \geq w_k > \dot{w}_k, \text{ and } p = p'\}$. Then, we can show that $T(\dot{w}, \ddot{w}, \dddot{w}; p')$ has uncountably many elements. Note that by conditions $iii)$ and $v)$, $\min\{\ddot{w}_k, \dddot{w}_k - \dot{w}_j\} > \dot{w}_k$. Let $a \in \mathbb{R}^4$ satisfy that $a_i = \frac{1-\dot{w}_j}{2}$, $a_j = \dot{w}_j$, $\min\{\ddot{w}_k, \dddot{w}_k - \dot{w}_j\} \geq a_k > \dot{w}_k$, and $a_y = 1 - a_i - a_j - a_k$. Since $1 > \dot{w}_j > 0$ and $\dot{w}_k > 0$, we have that $a_i, a_j, a_k \in (0, 1)$. Note that $a_y = 1 - \frac{1-\dot{w}_j}{2} - \dot{w}_j - a_k > \frac{1}{2} - \frac{\dot{w}_j}{2} - \dot{w}_k > \frac{1}{4}$, that is, $a_y \in (\frac{1}{4}, 1)$. Since $a_i + a_j + a_k + a_y = 1$, a is an allocation in the four-player model. Since a satisfies all requirements to be an allocation in $T(\dot{w}, \ddot{w}, \dddot{w}; p')$, we have that $(a, p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$. It is easily seen that for any $\varepsilon \in (0, a_k - \dot{w}_k)$, $((a_i, a_j, a_k - \varepsilon, a_y + \varepsilon), p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$.

We will prove that $T(\dot{w}, \ddot{w}, \dddot{w}; p')$ contains some state that S cannot contain or dominate. First, we show that there exists a state in $T(\dot{w}, \ddot{w}, \dddot{w}; p')$ that S cannot contain. Note that for any distinct states (w, p'), $(w', p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$, we have that $w_i = w'_i$, $w_j = w'_j$, and either $w'_y > w_y > w_k > w'_k$ or $w_y > w'_y >$

$w'_k > w_k$. If $w'_y > w_y > w_k > w'_k$, then (w', p') dominates (w', p') by player y pillaging player k. Similarly, if $w_y > w'_y > w'_k > w_k$, then (w, p') dominates (w', p'). Therefore, internal stability of S means that $T(\dot{w}, \ddot{w}, \dddot{w}; p') \cap S$ has at most one element. In addition, for any $(w, p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$, there exists a state $(w'', p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$ such that $w''_i = w_i$, $w''_j = w_j$, and $w''_y > w_y > w_k > w''_k$; that is, every state in $T(\dot{w}, \ddot{w}, \dddot{w}; p')$ is dominated by another state in $T(\dot{w}, \ddot{w}, \dddot{w}; p')$. Therefore, there exists a state $(w^T, p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$ such that (w^T, p') dominates a state in $T(\dot{w}, \ddot{w}, \dddot{w}; p') \cap S$, which can be empty. Then we have that $(w^T, p') \notin S$. Next, we show that S cannot dominate (w^T, p').

Let the set of states $T_1(w^T, p') = \{(w, p) : (w, p)$ results from player i pillaging player j at $(w^T, p')\}$. Note that $w^T_k + w^T_y = 1 - w^T_i - w^T_j = 1 - \frac{1-\dot{w}_j}{2} - \dot{w}_j = \frac{1-\dot{w}_j}{2} = w^T_i$. Let $(w^{T_1}, p) \in T_1(w^T, p')$. Since $w^T_i = w^T_k + w^T_y$, $w^T_j = \dot{w}_j$, and $w^T_k > \dot{w}_k$, we have that $w^{T_1}_i > w^{T_1}_k + w^{T_1}_y$, $w^{T_1}_j < w^T_j = \dot{w}_j$, and $w^{T_1}_k = w^T_k > \dot{w}_k$. Since $(\dot{w}, p') \in S \cap B^i_{j,k}(\frac{1}{2}, p')$, Lemma 29 implies that $[0, \dot{w}_j] \subset \{w_j : (w, p') \in S \cap B^i_{j,k}(\frac{1}{2}, p')\}$. Therefore, $S \cap B^i_{j,k}(\frac{1}{2}, p')$ contains $(\dot{w}^{T_1}, p')$ with $\dot{w}^{T_1}_j = w^{T_1}_j$. Since $\dot{w}_j > \dot{w}^{T_1}_j$, by Lemma 30, we have that $\dot{w}_k \geq \dot{w}^{T_1}_k$ and thus that $w^{T_1}_k > \dot{w}^{T_1}_k$. Therefore, (w^{T_1}, p) in $T_1(w^T, p')$ is dominated by $(\dot{w}^{T_1}, p')$ in $S \cap B^i_{j,k}(\frac{1}{2}, p')$ by player i pillaging players k and y. Since $(w^{T_1}, p) \in T_1(w^T, p')$ is arbitrary, every state in $T_1(w^T, p')$ is dominated by some state in $S \cap B^i_{j,k}(\frac{1}{2}, p')$ through player i pillaging players k and y. Therefore, we must have that $T_1(w^T, p') \cap S = \varnothing$.

Let the set of states $T_2(w^T, p') = \{(w, p) : (w, p)$ results from player i pillaging player k at $(w^T, p')\}$. Then for each $(w^{T_2}, p') \in T_2(w^T, p')$, we have that $w_i^{T_2} > w_k^{T_2} + w_y^{T_2}$, $w_j^{T_2} = \dot{w}_j$, and $w_y^{T_2} = w_y^T > w_k^T > \dot{w}_k$ since $w_i^T = w_k^T + w_y^T$, $w_j^T = \dot{w}_j$, and $w_y^T > \frac{1}{4} > w_k^T$. Therefore, every state in $T_2(w^T, p')$ is dominated by $(\dot{w}, p') \in S \cap B_{j,k}^i(\frac{1}{2}, p')$ through players i and k pillaging player y when $\dot{w}_k > w_k^{T_2}$, through player i pillaging player y when $\dot{w}_k = w_k^{T_2}$, or through player i pillaging players k and y when $\dot{w}_k < w_k^{T_2}$. Therefore, we have that $T_2(w^T, p') \cap S = \varnothing$.

Let the set of states $T_3(w^T, p') = \{(w, p) : (w, p)$ results from player i pillaging player y at $(w^T, p')\}$. Then for each $(w^{T_3}, p') \in T_3(w^T, p')$, we have that $w_i^{T_3} > w_k^{T_3} + w_y^{T_3}$, $w_j^{T_3} = \dot{w}_j$, and $w_k^{T_3} = w_k^T > \dot{w}_k$ since $w_i^T = w_k^T + w_y^T$, $w_j^T = \dot{w}_j$, and $w_k^T > \dot{w}_k$. Therefore, every state in $T_3(w^T, p')$ is dominated by $(\dot{w}, p') \in S \cap B_{j,k}^i(\frac{1}{2}, p')$ through player i either pillaging players k and y when $w_y^{T_3} > 0$, or pillaging player k when $w_y^{T_3} = 0$. Therefore, we have that $T_3(w^T, p') \cap S = \varnothing$.

Let the set of states $T_4(w^T, p') = \{(w, p) : (w, p)$ results from player y pillaging player k at $(w^T, p')\}$. Let $(w^{T_4}, p') \in T_4(w^T, p')$. Since $w_i^T = w_k^T + w_y^T$ and $w_y^T > \frac{1}{4} > \dot{w}_k$, we have that $w_i^{T_4} \geq w_y^{T_4} > \dot{w}_k$ and $w_j^{T_4} = \dot{w}_j$. Note that dominance relation in $T(\dot{w}, \ddot{w}, \dddot{w}; p')$ is transitive; that is, if $(w, p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$ is dominated by $(w', p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$ and (w', p') is dominated by $(w'', p') \in T(\dot{w}, \ddot{w}, \dddot{w}; p')$, then because $w_i = w_i' = w_i''$, $w_j = w_j' = w_j''$, and $w_y'' > w_y' > w_y > w_k > w_k' > w_k''$, (w, p') is dominated by (w'', p') through player y pillaging player k at (w, p'). Therefore, if

$w_k^{T_4} > \dot{w}_k$, then (w^{T_4}, p) is still in $T(\dot{w}, \ddot{w}, \dddot{w}; p')$ and thus since (w^T, p') dominates a state in $T(\dot{w}, \ddot{w}, \dddot{w}; p') \cap S$, which can be empty, (w^{T_4}, p') dominates a state in $T(\dot{w}, \ddot{w}, \dddot{w}; p') \cap S$. This means that $(w^{T_4}, p') \notin S$. If $\dot{w}_k \geq w_k^{T_4} > 0$, then (w^{T_4}, p') is dominated by $(\dot{w}, p')$ in $S \cap B^i_{j,k}(\frac{1}{2}, p')$ either through player i pillaging player y when $\dot{w}_k = w_k^{T_4}$, or through players i and k pillaging player y when $\dot{w}_k > w_k^{T_4}$. If $w_k^{T_4} = 0$, then $w_i^{T_4} = w_y^{T_4} > w_j^{T_4}$, and thus (w^{T_4}, p') is dominated by some state in the core such that players i and y have halves. Since $(w^{T_4}, p') \in T_4(w^T, p')$ is arbitrary, we have that $T_4(w^T, p') \cap S = \varnothing$.

Let the set of states $T_5(w^T, p') = \{(w, p) : (w, p)$ results from player y pillaging player j at $(w^T, p_1)\}$. Then for each $(w^{T_5}, p) \in T_5(w^T, p')$, we have that $w_i^{T_5} + w_k^{T_5} > w_j^{T_5} + w_y^{T_5}$ and $\frac{1}{2} > w_i^{T_5} > w_k^{T_5}$ since $w_i^T > w_y^T$, $w_k^T > w_j^T$ and $\frac{1}{2} > w_i^T > w_k^T$. Therefore, every state in $T_5(w^T, p')$ is dominated by some state in the core such that players i and k have halves. Therefore, we have that $T_5(w^T, p') \cap S = \varnothing$.

Let the set of states $T_6(w^T, p') = \{(w, p) : (w, p)$ results from player k pillaging player j at $(w^T, p')\}$. Then for each $(w^{T_6}, p) \in T_6(w^T, p')$, we have that $w_i^{T_6} + w_y^{T_6} > w_j^{T_6} + w_k^{T_6}$ and $\frac{1}{2} > w_i^{T_6} > w_y^{T_6}$ since $\frac{1}{2} > w_i^T > w_y^T > w_k^T > w_j^T$. Therefore, every state in $T_6(w^T, p')$ is dominated by some state in the core such that players i and y have halves. Therefore, we have that $T_6(w^T, p') \cap S = \varnothing$.

Let the set of states $T_7(w^T, p') = \{(w, p) : (w, p)$ results from players i and y pillaging player k at $(w^T, p')\}$. Then for each $(w^{T_7}, p) \in T_7(w^T, p')$, we have that $w_i^{T_7} > w_k^{T_7} + w_y^{T_7}$, $w_j^T = \dot{w}_j$, and $w_y^{T_7} > \dot{w}_k$ since $w_i^T = w_k^T + w_y^T$ and $w_y^T > \dot{w}_k$. Therefore, every state in $T_7(w^T, p')$ is dominated by $(\dot{w}, p') \in S \cap B^i_{j,k}(\frac{1}{2}, p')$. Therefore, we have that $T_7(w^T, p') \cap S = \varnothing$.

Let the set of states $T_8(w^T, p') = \{(w, p) : (w, p)$ results from players i and y pillaging player j at $(w^T, p')\}$. Let $(w^{T_8}, p'') \in T_8(w^T, p')$. Since $w_i^T > w_j^T + w_y^T$, $\ddot{w}_k \geq w_k^T$, and $w_y^T > \frac{1}{4} > \ddot{w}_j$, we have that $w_i^{T_8} > w_j^{T_8} + w_y^{T_8}$, $\ddot{w}_k \geq w_k^T = w_k^{T_8}$, and $w_y^{T_8} > w_y^T > \ddot{w}_j$. When $w_k^{T_8} = \ddot{w}_k$, (w^{T_8}, p'') is dominated by $(\ddot{w}, p'') \in S \cap B^i_{k,j}(\frac{1}{2}, p'')$ through player i pillaging players j and y when $\ddot{w}_j < w_j^{T_8}$, through player i pillaging player y when $\ddot{w}_j = w_j^{T_8}$, or through players i and j pillaging player y when $\ddot{w}_j > w_j^{T_8}$. Now, we check the case that $\ddot{w}_k > w_k^{T_8}$. Since $(\ddot{w}, p'') \in S \cap B^i_{k,j}(\frac{1}{2}, p'')$, Lemma 29 implies that $[0, \ddot{w}_k] \subset \{w_k : (w, p'') \in S \cap B^i_{k,j}(\frac{1}{2}, p'')\}$. Therefore, $S \cap B^i_{k,j}(\frac{1}{2}, p'')$ contains the state $(\ddot{w}^{T_8}, p'')$ such that $\ddot{w}_k^{T_8} = w_k^{T_8}$. Since $\ddot{w}_k > \ddot{w}_k^{T_8}$, Lemma 30 implies that $\ddot{w}_j \geq \ddot{w}_j^{T_8}$ and thus that $w_y^{T_8} > w_y^T > \frac{1}{4} > \ddot{w}_j \geq \ddot{w}_j^{T_8}$. Therefore, $(w^{T_8}, p'') \in T_8(w^T, p')$ is dominated by the state $(\ddot{w}^{T_8}, p'') \in S \cap B^i_{k,j}(\frac{1}{2}, p'')$ through player i pillaging players j and y when $\ddot{w}_j^{T_8} < w_j^{T_8}$, through player i pillaging player y when $\ddot{w}_j^{T_8} = w_j^{T_8}$, or through players i and j pillaging player y when $\ddot{w}_j^{T_8} > w_j^{T_8}$. Since $(w^{T_8}, p'') \in T_8(w^T, p')$ is arbitrary, we have that $T_8(w^T, p') \cap S = \varnothing$.

Let the set of states $T_9(w^T, p') = \{(w,p) : (w,p)$ results from players i and k pillaging player y at $(w^T, p')\}$. Then for each $(w^{T_9}, p') \in T_9(w^T, p')$, we have that $w_i^{T_9} > w_k^{T_9} + w_y^{T_9}$, $w_j^{T_9} = \dot{w}_j$, and $w_k^{T_9} > \dot{w}_k$ since $w_i^T = w_k^T + w_y^T$ and $w_k^T > \dot{w}_k$. Therefore, every state in $T_9(w^T, p')$ is dominated by $(\dot{w}, p') \in S \cap B^i_{j,k}(\frac{1}{2}, p')$. Therefore, we have that $T_9(w^T, p') \cap S = \varnothing$.

Let the set of states $T_{10}(w^T, p') = \{(w,p) : (w,p)$ results from players i and k pillaging player j at $(w^T, p')\}$. Let $(w^{T_{10}}, p) \in T_{10}(w^T, p')$. Since $w_i^T > w_y^T > \frac{1}{4}$, $\mathring{w}_j \geq \mathring{w}_k = \dot{w}_j = w_j^T$, and $\dddot{w}_k - \dot{w}_j \geq w_k^T > \dot{w}_j$, we have that $w_i^{T_{10}} > w_y^{T_{10}} > \frac{1}{4}$, $\mathring{w}_j \geq w_j^T > w_j^{T_{10}}$, and $\dddot{w}_k > w_k^{T_{10}} > \mathring{w}_k$. Since $(\dddot{w}, p''') \in S \cap B^i_{k,j}(\frac{1}{2}, p''')$, Lemma 29 implies that $[0, \dddot{w}_k] \subset \{w_k : (w, p''') \in S \cap B^i_{k,j}(\frac{1}{2}, p''')\}$. Therefore, $S \cap B^i_{k,j}(\frac{1}{2}, p'')$ contains the state $(\dddot{w}^{T_{10}}, p''')$ with $\dddot{w}_k^{T_{10}} = w_k^{T_{10}}$. Since $\dddot{w}_k > w_k^{T_{10}} > \mathring{w}_k$, Lemma 30 means that $\dddot{w}_j \geq \dddot{w}_j^{T_{10}} \geq \mathring{w}_j$. Since $w_y^{T_{10}} > \frac{1}{4} > \dddot{w}_j \geq \dddot{w}_j^{T_{10}} > w_j^{T_{10}}$, $(\dddot{w}^{T_{10}}, p''') \in S \cap B^i_{k,j}(\frac{1}{2}, p'')$ dominates the state $(w^{T_{10}}, p)$ by players i and j pillaging player y. Since $(w^{T_{10}}, p) \in T_{10}(w^T, p')$ is arbitrary, every state in $T_{10}(w^T, p')$ is dominated by some state in $S \cap B^i_{k,j}(\frac{1}{2}, p''')$. Therefore, we have that $T_{10}(w^T, p') \cap S = \varnothing$.

Let the set of states $T_{11}(w^T, p') = \{(w,p) : (w,p)$ results from players k and y pillaging player j at $(w^T, p')\}$. Then for each $(w^{T_{11}}, p) \in T_{11}(w^T, p')$, we have that $w_i^{T_{11}} < w_j^{T_{11}} + w_k^{T_{11}} + w_y^{T_{11}}$, $\frac{1}{2} > w_y^{T_{11}}$, and $\frac{1}{4} > w_k^{T_{11}} > w_j^{T_{11}}$ since $w_i^T < w_j^T + w_k^T + w_y^T$, $\frac{1}{2} > w_y^T + w_j^T$, and $\frac{1}{4} > \dddot{w}_k \geq w_k^T + w_j^T$. Note that by Lemma 26, $D \subset S$. Therefore, every state in $T_{11}(w^T, p')$ is dominated by some state (w, p) in D such that $w_y = \frac{1}{2}$, $w_j = w_k = \frac{1}{4}$, and $p^{p'_i} = I$ through players j, k, and y pillaging player i. Therefore, we have that $T_{11}(w^T, p') \cap S = \varnothing$.

Finally, let the set of states $T_{12}(w^T, p') = \{(w, p) : (w, p)$ results from players i, k, and y pillaging player j at $(w^T, p')\}$. Let $(w^{T_{12}}, p) \in T_{12}(w^T, p')$. Since $w_i^T > w_y^T + w_j^T$ and $w_i^T > w_y^T > \frac{1}{4} > w_k^T + w_j^T$, we have that $w_i^{T_{12}} > w_y^{T_{12}} > \frac{1}{4} > w_j^{T_{12}} + w_k^{T_{12}}$. If $w_i^{T_{12}} > \frac{1}{2}$, then $(w^{T_{12}}, p)$ is dominated by some state in the core such that player i has the total wealth. If $w_i^{T_{12}} \leq \frac{1}{2}$, then $(w^{T_{12}}, p)$ is dominated by some state in the core at which players i and j have halves. Since $(w^{T_{12}}, p') \in T_{12}(w^T, p')$ is arbitrary, we have that $T_{12}(w^T, p') \cap S = \varnothing$.

Therefore, $(w^T, p) \notin S$ cannot be dominated by any state in S. This contradiction shows that there is no stable set in the four-player and two-region model. $\square$

Theorem 35 generalizes Proposition 34 to a $I-$player and $N-$region model where $I = 4$ and $N = 2$, or $I \geq 4$ and $N \geq 3$.

Theorem 35. *No stable set exists in a $I-$player and $N-$region model where $I = 4$ and $N = 2$, or $I \geq 4$ and $N \geq 3$.*

Proof. Suppose by way of contradiction that there exists a stable set S. For any four distinct players i, j, k, and y, define the set $F(i, j, k, y)$ of states by $F(i, j, k, y) = \{(w, p) : w_i + w_j + w_k + w_y = 1,\ p^1 \cup p^2 = \{i, j, k, y\}$, and $p^3 =$

$I \setminus \{i, j, k, y\}$ if $N \geq 3\}$. Then any state in $F(i, j, k, y)$ cannot be dominated by another state in $X \setminus F(i, j, k, y)$. Thus $S \cap F(i, j, k, y)$ is externally stable with respect to $F(i, j, k, y)$. Obviously $S \cap F(i, j, k, y)$ is internally stable. Therefore, $S \cap F(i, j, k, y)$ is both internal stable and external stable with respect to $F(i, j, k, y)$. It is easily seen that $S \cap F(i, j, k, y)$ of states can be adapted for a stable set in the four-player and two-region model. This contradicts Proposition 34, which shows nonexistence of stable set in the four-player and two-region model. This contradiction completes the proof. □

2.3. Core in expectation

As has been shown in section 2.2, stable set is not appropriate for a solution to the spatial pillage game. In a $I-$player and $N-$region model where $I = 4$ and $N = 2$, or $I \geq 4$ and $N \geq 3$, no stable set exists. In three-player models, there exist stable sets. However, they contain implausible states, such as some states in the set of states $X_{\#I} = \{(w, p) :$ for some play i, $1 > w_i > \frac{1}{2}\}$. Every state in $X_{\#I}$ is naturally thought not to be stationary because one of the players has enough power to pillage the others, so the player would achieve all of the wealth.

These problems with the solution concept of stable set are caused by weakness of external stability, which requires that any state outside a stable set be directly dominated by some state in the stable set. Some states in $X_{\#I}$ are directly dominated only by other states in $X_{\#I}$, thus a stable set must contain some states in $X_{\#I}$ to satisfy external stability. This shows why stable sets in three-player models contain some states in $X_{\#I}$. In $I-$player and $N-$region models where $I = 4$ and $N = 2$, or $I \geq 4$ and $N \geq 3$, if an internally stable set S' includes a set of states that dominates every state in $X_{\#I} \backslash S'$, then there exists some state $(w, p) \notin S'$ that S' cannot dominate. It is because the internally stable set S' inevitably dominates every state that dominates (w, p). This shows why no stable set exists in these models.

Jordan (2005) introduced the solution concept of **farsighted core** of **consistent expectation** that is based on the concept of **dominance in expectation**. We can settle the problem with the solution concept of stable set by adopting the solution concept of farsighted core, as shown in Theorem 40 and Lemma 41. Theorem 40 assures that a farsighted core of a consistent expectation exists. Lemma 41 guarantees that a farsighted core of a consistent expectation does not contains any state in $X_{\#I}$. Note that the definitions of expectation and dominance in expectation are adapted for a spatial pillage game.

An **expectation** is a belief that all players have in common and indicates how each state proceeds.

Definition 36. *An **expectation** is a function $f : X \longrightarrow X$ satisfying, for some integer $k \geq 2$, $f^k = f^{k-1}$ where $f^k = f \circ f^{k-1}$. Let $f_w(w,p)$ and $f_p(w,p)$ denote the allocation and the distribution at $f(w,p)$, respectively.*

Dominance in Expectation between states indicates the possible state that the present state can change to provided that players follow the expectation after a change. Both *physical* and *spatial conditions* should be satisfied in order for a *winning coalition in expectation*, who end up being better off, to change its present state through defeating a *losing coalition in expectation*, who end up being worse off. Physical and spatial conditions are reflected on condition *iii*) and conditions *i*) and *ii*) in Definition 37, respectively.

Definition 37. *Let an expectation f satisfy that $f^k = f^{k+1}$. Given states (w,p) and (w',p'), define $W_f = \{i : f_w^k(w',p')_i > w_i\}$ and $L_f = \{i : f_w^k(w',p')_i < w_i\}$. Suppose that for some $r \in R$, i) $\{i : w'_i \neq w_i\} \subset p'^r$; ii) for all $q \neq r$, $p'^q = p^q \setminus (W_f \cap p'^r)$; and iii) $\sum_{i \in W_f \cap p'^r} w_i > \sum_{i \in L_f \cap p'^r} w_i$. Then (w',p') **dominates** (w,p) **in expectation**.*

An expectation is **consistent** if it is organized in accord with the relation of dominance in expectation. If a state (w,p) proceeds to another state (w',p') in expectation, then (w',p') dominates (w,p) in expectation. If a state (w'',p'') is stationary in expectation, then no state dominates (w'',p'') in expectation.

Definition 38. *An expectation f is **consistent** if $f(w,p)$ dominates (w,p) in expectation when $f(w,p) \neq (w,p)$ and (w,p) is undominated in expectation when $f(w,p) = (w,p)$.*

Farsighted core and **farsighted supercore**[1] are defined as follows.

Definition 39. *Given an expectation f, the **farsighted core** under the expectation f is the set of states $K_f = \{(w,p) \in X :$ under the expectation f, no state in X dominates (w,p) in expectation$\}$. The **farsighted supercore** C_S is the intersection of all farsighted cores of consistent expectations.*

Theorem 40 assures that there exists a consistent expectation that has the set D of states as farsighted core.

Theorem 40 (Existence of a consistent expectation)**.** *There exists a consistent expectation f such that $K_f = D$.*

Proof. Let $\bar{X} = \{(w,p) \in X :$ for each region $r \in R$, $\sum_{i \in p^r} w_i = 0, \frac{1}{2}$, or 1 and for some region $q \in R$, if $\sum_{i \in p^q} w_i = \frac{1}{2}$, then for some player z, $w_z = \frac{1}{2}\}$. Then by Lemma 26, D is the unique set that satisfies both internal stability and external stability with respect to $\bar{X}$. Therefore, it suffices to construct a consistent expectation f such that for some positive integer k, $f^k(X \setminus \bar{X}) \subset D$.

[1]***Farsighted supercore* is named after Roth's (1976) *supercore*.**

For any positive integer $n \geq 2$, define $X^n = \{(w,p)$: there exist n distinct regions $r_1, ..., r_n$ such that $i)$ $\sum_{i \in p^{r_1} \cup \cdots \cup p^{r_n}} w_i = 1$; $ii)$ $\sum_{i \in p^{r_1}} w_i = \max_{r' \in R}\{\sum_{i \in p^{r'}} w_i\} > \sum_{i \in p^{r_n}} w_i > 0$; and $iii)$ for each $j \in p^{r_1}$ and some nonnegative integer k_j, $w_j = (\sum_{z \in p^{r_1}} w_z) \times (\frac{1}{2})^{k_j}\}$. Then any state in X^2 is dominated by some state in D through players in a wealthier region pillaging other players in another region. Therefore, we construct f such that every state in X^2 is dominated in expectation by some state in D. Similarly, we can construct f such that for any integer $k \geq 2$, any state in $X^{k+1} \setminus X^k$ is dominated in expectation by some state in X^k. Note that for any state $(w,p) \in X^n$, if $f_w(w,p)_i > w_i$, then for any positive integer k, $f^k_w(w,p)_i \geq f_w(w,p)_i$ and that during the change from (w,p) to $f^{n-1}(w,p)$, pillage happens in each region at most once.

Now, we design the expectation f to satisfy that any state in $X \setminus (X^{\#R} \cup \bar{X})$ is dominated in expectation by some state in $X^{\#R}$. For some regions q and q', let $(w,p) \in X \setminus X^{\#R}$ satisfy that $\sum_{i \in p^q} w_i > \sum_{i \in p^{q'}} w_i > 0$. Then we have that $(w,p) \notin \bar{X}$. Since $(w,p) \notin X^{\#R}$, there exist region r and player i such that $i \in p^r$, $\sum_{z \in p^r} w_z = \max_{r' \in R}\{\sum_{z \in p^{r'}} w_z\}$, and for any nonnegative integer k_i, $w_i \neq (\sum_{z \in p^r} w_z) \times (\frac{1}{2})^{k_i}$. Theorem 25 assures that there exists the state $(w',p) \in X^{\#R}$ such that $i)$ $w'_i = w_i$ when $i \notin p^r$; $ii)$ for some nonnegative integer k_i, $w'_i = (\sum_{z \in p^r} w_z) \times (\frac{1}{2})^{k_i}$ when $i \in p^r$; and $iii)$ $\sum_{z \in \{y : w'_y > w_y\}} w_z > \sum_{z \in \{y : w'_y < w_y\}} w_z$.

Then $(w', p) \in X^{\#R}$ dominates (w, p), and thus we can make (w', p) dominate (w, p) in expectation.

Let $(\dot{w}, p) \in X \setminus \bar{X}$ satisfy that for any two regions r and q, if $\sum_{i \in p^r} \dot{w}_i > 0$ and $\sum_{i \in p^q} \dot{w}_i > 0$, then $\sum_{i \in p^r} \dot{w}_i = \sum_{i \in p^q} \dot{w}_i$. Note that $(\dot{w}, p) \notin X^{\#R}$. For some distinct regions r and q, if $\sum_{i \in p^r} \dot{w}_i = \sum_{i \in p^q} \dot{w}_i = \frac{1}{2}$ and $\dot{w}$ is dyadic, then since $(\dot{w}, p) \notin \bar{X}$, for each i, $\dot{w}_i < \frac{1}{2}$. In this case, a coalition E such that $\sum_{i \in E} \dot{w}_i = \frac{1}{2}$, $p^r \nsubseteq E$ (or $p^q \nsubseteq E$), and $E \subset \{i : w_i > 0\}$ can pillage all players in one of the regions and divide their booty in proportion to their wealth. For each region r, if $\sum_{i \in p^r} \dot{w}_i = 0$ or $\frac{1}{2}$ and there exists player i such that for any nonnegative integer k_i, $\dot{w}_i \neq (\frac{1}{2})^{k_i}$, then by Theorem 25, there exists an allocation $\dot{w}'$ such that $i)$ when $z \notin p^{p_i}$, $\dot{w}'_z = \dot{w}_z$; $ii)$ when $z \in p^{p_i}$, for some nonnegative integer k_z, $\dot{w}'_z = (\sum_{y \in p^{p_i}} \dot{w}_y) \times (\frac{1}{2})^{k_z}$; and $iii)$ $\sum_{z \in \{y : \dot{w}'_y > \dot{w}_y\}} \dot{w}_z > \sum_{z \in \{y : \dot{w}'_y < \dot{w}_y\}} \dot{w}_z$. Then a coalition $\dot{E}$ that consists of $\{z : \dot{w}'_z > \dot{w}_z\}$ and player j such that $p_j \neq p_i$ and $\dot{w}_j > 0$ can pillage the other players at region p_i and proportion their wealth to $\dot{w}'$ while giving $\frac{1}{2}(\frac{1}{2} + \dot{w}_j)$ to player j. Note that players in the coalition $\dot{E}$ who earn nothing or even lose their wealth at $(\dot{w}', p')$ participate in $\dot{E}$ because they expect that their wealth will increase in the future movement. Similarly, in case that for each region r, $\sum_{i \in p^r} \dot{w}_i < \frac{1}{2}$, we can construct f so that $(\dot{w}, p)$ is dominated in expectation by some state in $X^{\#R}$. Note that for any $(w, p) \in X \setminus (X^{\#R} \cup \bar{X})$, if $f_w(w, p)_i > w_i$, then for any positive integer k, $f_w^k(w, p)_i \geq f_w(w, p)_i$ and that during the change from (w, p) to $f^{\#R}(w, p)$, pillage happens in each region at most once.

Now, we only need to check if the expectation f is consistent. The expectation f satisfies that $i)$ for any $(w, p) \in X$, if $f_w(w, p)_i > w_i$, then for any positive integer k, $f_w^k(w, p)_i \geq f_w(w, p)_i$; and $ii)$ during the change from (w, p) to $f^{\#R}(w, p)$, pillage can happen in each region at most once. If a player experiences pillage in his region, then he will never be pillaged during the rest of the process of f; that is, for any $(w, p) \in X \setminus K_f$ and any $i \in I$, if p_i satisfies that $\{z : f_w(w, p)_z \neq w_z\} \subset p^{p_i}$, then any positive integer k, $f_w^k(w, p)_i \geq f_w(w, p)_i$. It is easily seen that for any state $(w, p) \in X$ and any positive integer k, if $f^k(w, p) \neq f^{k-1}(w, p)$, then $i)$ there exists a region $r \in R$ such that $\{z : f_w^k(w, p)_z \neq f_w^{k-1}(w, p)_z\} \subset f_p^k(w, p)^r$; $ii)$ for all $q \neq r$, $f_p^k(w, p)^q = f_p^{k-1}(w, p)^q \setminus \{z : f_w^k(w, p)_z > f_w^{k-1}(w, p)_z\}$; and $iii)$ $\sum_{y \in \{z : f_w^k(w,p)_z > f_w^{k-1}(w,p)_z\}} f_w^{k-1}(w, p)_y > \sum_{y \in \{z : f_w^k(w,p)_z > f_w^{k-1}(w,p)_z\}} f_w^{k-1}(w, p)_y$. These show that $f(w, p)$ dominates (w, p) in expectation when $f(w, p) \neq (w, p)$ and (w, p) is undominated in expectation when $f(w, p) = (w, p)$. Therefore, f is consistent.

$\square$

The expectation f constructed above shows how the neutrality assumption is modified in a spatial pillage game. If some players expect that they would be pillaged during the process of f, then they combine their power under spatial restriction to protect themselves although some of them are not pillaged immediately. For example, the state $((\frac{1}{2}, \frac{1}{6}, \frac{1}{3}), (1, 2, 2))$ is not dominated in expectation by $((\frac{2}{3}, 0, \frac{1}{3}), (2, 2, 2))$ because player 3 will be against player 1 to protect player 2 in the expectation that after player 1 pillaging player 2, player 1 would pillage player 3. This shows that neutrality is modified. However, $((\frac{1}{2}, \frac{1}{6}, \frac{1}{3}), (1, 2, 2))$ is dominated in expectation by $((\frac{1}{2}, 0, \frac{1}{2}), (1, 2, 2))$ because player 1 keeps neutral. Also, the state $((\frac{1}{2}, \frac{1}{6}, \frac{1}{3}), (1, 1, 2))$ is dominated in expectation by $((\frac{2}{3}, 0, \frac{1}{3}), (1, 1, 2))$ because player 3 cannot protect player 2 because of spatial restriction. In these cases, neutrality is not modified. Therefore, as Jordan (2005) said, "the concept of domination in expectation constitutes an endogenous modification of the neutrality assumption" in a spatial pillage game.

Lemma 41 shows that in a consistent expectation, if one of players has a majority of the wealth, then the player would finally have all of the wealth. Lemmas 41 and 42 are used to prove Proposition 43.

Lemma 41. *Let the set of states* $X_{\#I} = \{(w, p) : \textit{for some play } i,\ 1 > w_i > \frac{1}{2}\}$ *and the set of states* $D_0 = \{(w, p) : \textit{for each } i,\ w_i = 0 \textit{ or } 1\}$*. Then for any consistent expectation* f*, there exists a positive integer* k *such that* $f^k(X_{\#I}) = D_0$*.*

Proof. For any integer $n \geq 2$, define $X_n = \{(w, p) :$ there exist n distinct players $i_1, ..., i_n$ such that $1 > w_{i_1} > \frac{1}{2}$ and $w_{i_1} + \cdots + w_{i_n} = 1\}$. Let f be a consistent expectation such that $f^k = f^{k+1}$. Then we have that $X_2 \cap K_f = \varnothing$ because every state in X_2 is dominated in expectation by some state in $D_0 \subset K_f$. Suppose that for some integer $n \geq 2$, $X_n \cap K_f = \varnothing$, then we can show that $X_{n+1} \cap K_f = \varnothing$. For two distinct players i and j, let $(w', p') \in X_{n+1}$ and (w, p) satisfy that $w'_i > \frac{1}{2}$, $w'_j > 0$, $w_i = w'_i + w'_j$, $w_z = w'_z$ and $p_z = p'_z$ for each $z \notin \{i, j\}$, and $p_i = p_j = p'_j$. Then the state (w, p) is in X_n, and thus K_f does not contain (w, p). Since $w_i > \frac{1}{2}$, $f^k(w, p)$ satisfies that $f^k_w(w, p)_i \geq w_i$. Otherwise any change in (w, p) is not possible. Since $w_i > w'_i$, the state (w, p) dominates (w', p') in expectation, and so K_f does not contain (w', p'). Since (w', p') is arbitrary, we have that $X_{n+1} \cap K_f = \varnothing$. Consequently, we have that $X_{\#I} \cap K_f = \varnothing$. It is easily seen that if $(w'', p'') \in X_{\#I}$ and $w''_i > \frac{1}{2}$, then $f^k_w(w'', p'')_i \geq w''_i$. Therefore, we have that $f^k(X_{\#I}) = D_0$. □

Lemma 42 (Lemma 3.10 in Jordan, 1999). *For some positive integer k, let w be a dyadic allocation such that for each i, if $w_i > 0$ then $w_i \geq 2^{-(k+1)}$. If an allocation w' satisfies that $\sum_{z \in \{i : w'_i > w_i\}} w_z > \sum_{z \in \{i : w'_i < w_i\}} w_z$, then there exists a dyadic allocation w'' such that $\sum_{z \in \{i : w''_i > w'_i\}} w'_z > \sum_{z \in \{i : w''_i < w'_i\}} w'_z$ and for each i, if $w''_i > 0$ then $w''_i \geq 2^k$.*

Proposition 43 shows that every farsighted core of a consistent expectation includes the set D^* of dyadic states at which one of players has a half of the wealth or the total wealth.

Proposition 43. *Let $D^* = \{(w,p) \in D :$ there exists player i with $w_i = 1$ or $\frac{1}{2}\}$. Then the farsighted supercore C_s includes D^*.*

Proof. [2]For any player i, define the set $X(i)$ of states by $X(i) = \{(w,p) \in X : w_i = \frac{1}{2}$ and for some region r, $\sum_{z \in p^r \setminus \{i\}} w_z = \frac{1}{2}\}$. Let f be a consistent expectation with $f^k = f^{k+1}$. We first show that $f^k(X(i)) \subset X(i)$. Let $(\dot{w}, \dot{p}) \in X(i)$. Note that only a winning coalition in expectation, $\{z : f_w^k(\dot{w}, \dot{p})_z > \dot{w}_z\}$, can emigrate to another region, that is, for some player z, if $f_p^k(\dot{w}, \dot{p})_z \neq \dot{p}_z$, then $f_w^k(\dot{w}, \dot{p})_z > \dot{w}_z$. By way of contradiction, suppose that $f_w^k(\dot{w}, \dot{p})_i > \frac{1}{2}$. Then by Lemma 41, $f^k(\dot{w}; \dot{p})_i = 1$. In this case, we have that $\sum_{z \in \{y : f_w^k(\dot{w}, \dot{p})_y > \dot{w}_y\}} \dot{w}_z = \sum_{z \in \{y : f_w^k(\dot{w}, \dot{p})_y < \dot{w}_y\}} \dot{w}_z$. Thus $f^k(\dot{w}, \dot{p})$ cannot dominate $(\dot{w}, \dot{p})$ in expectation because all players in the losing coalition in expectation, $\{z : f_w^k(\dot{w}, \dot{p})_z < \dot{w}_z\}$, are in a common region. Since $f_w^k(\dot{w}, \dot{p})_i < \frac{1}{2}$ is not possible, we must have that $f_w^k(\dot{w}, \dot{p})_i = \frac{1}{2}$. If there exists a region r such that $\sum_{z \in \dot{p}^r} f_w^k(\dot{w}, \dot{p})_z < \frac{1}{2}$, then $f_w^k(\dot{w}, \dot{p})$ is dominated in expectation by some state in $X_{\#I} = \{(w,p) :$ for some play z, $1 > w_z > \frac{1}{2}\}$ through player i pillaging players at region r. Therefore, for some region r, we have that $\sum_{z \in f_p^k(\dot{w}, \dot{p})^r \setminus \{i\}} f_w^k(\dot{w}, \dot{p})_z = \frac{1}{2}$ and thus that $f^k(\dot{w}, \dot{p}) \in X(i)$. Since $(\dot{w}, \dot{p}) \in X(i)$ is arbitrary, we have that $f^k(X(i)) \subset X(i)$.

[2]The proof of Proposition 43 is similar to the proof of Theorem in Jordan (2005).

Next, we show that $D^* \subset K_f$ and thus that $D^* \subset C_s$. Suppose by way of contradiction that $D^* \cap X(i) \not\subseteq f^k(X(i))$. For any nonnegative integer n, define $D_n = \{(w,p) \in D : \text{for each } i,\ w_i = 0 \text{ or } w_i \geq 2^{-n}\}$. Note that $\bigcup_{n\in\mathbb{N}} D_n = D^*$. Thus there exists a positive integer m with $m = \min\{n : D_n \cap X(i) \not\subseteq f^k(X(i))\}$. Since $D_1 \subset C_S$, m is greater than 1. Let $(\dot{w},\dot{p})$ be in $(D_m \cap X(i)) \setminus f^k(X(i))$. Since $(\dot{w},\dot{p}) \notin f^k(X(i))$, we have that $(\dot{w},\dot{p}) \neq f^k(\dot{w},\dot{p})$. Since f is consistent and $(\dot{w},\dot{p}), f^k(\dot{w},\dot{p}) \in X(i)$, we have that $\sum_{z\in\{y\neq i: f_w^k(\dot{w},\dot{p})_y > \dot{w}_y\}} \dot{w}_z > \sum_{z\in\{y\neq i: f_w^k(\dot{w},\dot{p})_y < \dot{w}_y\}} \dot{w}_z$ and thus that $\sum_{z\in\{y\neq i: 2f_w^k(\dot{w},\dot{p})_y > 2\dot{w}_y\}} 2\dot{w}_z > \sum_{z\in\{y\neq i: 2f_w^k(\dot{w},\dot{p})_y < 2\dot{w}_y\}} 2\dot{w}_z$. Note that the allocation $\mathring{w}$ that consists of players $I \setminus \{i\}$ such that for each $z \neq i$, $\mathring{w}_z = 2\dot{w}_z$ is a dyadic allocation in $(\#I - 1)-$player model. Therefore, by Lemma 42, there exists a dyadic allocation $\dot{w}'$ in $\#I-$player model such that $\dot{w}'_i = \frac{1}{2}$, $\sum_{z\in\{y\neq i: 2\dot{w}'_y > 2f_w^k(\dot{w},\dot{p})_y\}} 2f_w^k(\dot{w},\dot{p})_z > \sum_{z\in\{y\neq i: 2\dot{w}'_y < 2f_w^k(\dot{w},\dot{p})_y\}} 2f_w^k(\dot{w},\dot{p})_z$, and for each z, if $\dot{w}'_z > 0$ then $\dot{w}'_z \geq 2^{-(m-1)}$. Since $f^k(\dot{w},\dot{p}) \in X(i)$, there exists some region r such that $\sum_{z\in f_p^k(\dot{w},\dot{p})^r\setminus\{i\}} f_w^k(\dot{w},\dot{p})_z = \frac{1}{2}$. Let the distribution $\dot{p}'$ satisfy that for each z, if $\dot{w}'_z \neq f_w^k(\dot{w},\dot{p})_z$ then $\dot{p}'_z = r$, otherwise $\dot{p}'_z = f_p^k(\dot{w},\dot{p})_z$. If $(\dot{w}',\dot{p}') \in f^k(X(i))$, then $(\dot{w}',\dot{p}')$ dominates $f^k(\dot{w},\dot{p})$ in expectation. Therefore, we have that $(\dot{w}',\dot{p}') \notin f^k(X(i))$ and $(\dot{w}',\dot{p}') \in D_{m-1} \cap X(i)$. This contradicts the definition of m. Consequently, we have that for each i, $D^* \cap X(i) \subset f^k(X(i))$. Since $D_0 = (D^* \setminus X(i)) \subset K_f$, we have that $D^* \subset K_f$. Since f is an arbitrary consistent expectation, we have that $D^* \subset C_s$. □

Theorem 44 shows that D is the unique farsighted core in a $I-$player and $N-$region model where $1 \leq I \leq 3$ or $N = 1$. Note that Jordan (2005) used *one-step expectation,* where every state reaches its stationary state within one step, and had the same result as Theorem 44 in one-region models. Therefore, the definition of expectation in Jordan (2005) can be generalized to *finite-step expectation,* where some states take finite steps, possibly more than one step, to reach their stationary states.

Theorem 44. *In a $I-$player and $N-$region model where $1 \leq I \leq 3$ or $N = 1$, D is the unique farsighted core of consistent expectations, and thus the farsighted supercore is D.*

Proof. By Theorem 40, D is a farsighted core of some consistent expectation. To show uniqueness, we assume that f is a consistent expectation. We first prove that $D = K_f$ in one-region models. For any state $(w,p) \notin K_f$, we have that $\sum_{z\in\{i:f_w^k(w,p)_i>w_i\}} w_z > \sum_{z\in\{i:f_w^k(w,p)_i<w_i\}} w_z$. For any states $(w',p'),(w'',p'') \in K_f$, we have that $\sum_{z\in\{i:w_i''>w_i'\}} w_z' \leq \sum_{z\in\{i:w_i''<w_i'\}} w_z'$. Therefore, K_f satisfies external stability and internal stability, that is, K_f is a stable set. Theorem 25 implies that $K_f = D$.

Let $1 \leq I \leq 3$. Then for any $(w,p) \in D$, there exists player i with $w_i = 1$ or $\frac{1}{2}$. Thus we have that $(w,p) \in D^* = \{(w,p) \in D :$ there exists player i with $w_i = 1$ or

$\frac{1}{2}\}$. By Lemma 41, we have that $D = D^* \subset C_s \subset K_f$. To show that $K_f \subset D$, let $(w, p) \notin D$. If for some player i, $w_i > \frac{1}{2}$, then by Lemma 41, $f^k(w, p) \in D_0 \subset D$, and thus $(w, p) \notin K_f$. If for each player i, $w_i \leq \frac{1}{2}$, then (w, p) is dominated in expectation by some state in $D \subset K_f$ such that two players have halves through two players or one player pillaging another player. Thus we have that $(w, p) \notin K_f$ since f is consistent. Therefore, we have that $K_f = D$.

Since f is an arbitrary consistent expectation, D is the unique farsighted core. Therefore, the farsighted supercore is D in a $I-$player and $N-$region model where $1 \leq I \leq 3$ or $N = 1$. □

Example 45 provides a consistent expectation f with $D \nsubseteq K_f$ and $K_f \nsubseteq D$, that is, a farsighted core of a consistent expectation might contain nondyadic states and rule out dyadic states. Therefore, Example 45 shows that Theorem 44 cannot be generalized.

Example 45. *In the five-player three-region model, there exists a consistent expectation f such that for some $\varepsilon \in (0, \frac{1}{12})$, $K_f = D \cup \{((\frac{1}{4}+\varepsilon, \frac{1}{4}+\varepsilon, \frac{1}{4}+\varepsilon, 0, \frac{1}{4}-3\varepsilon), (1, 1, 2, 1, 3))\} \setminus \{((\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0), (1, 1, 1, 1, 2))\}$.*

Proof. Let $(w', p') = ((\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, 0), (1, 1, 1, 1, 2))$. First, construct f such that $(w', p') \longrightarrow f(w', p') = ((3\varepsilon, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}-3\varepsilon), (3, 1, 1, 1, 3)) \longrightarrow f^2(w', p') = ((2\varepsilon, \frac{1}{4}, \frac{1}{4}+\varepsilon, \frac{1}{4}, \frac{1}{4}-3\varepsilon), (2, 1, 2, 1, 3)) \longrightarrow f^3(w', p') = ((\frac{1}{4}+\varepsilon, \frac{1}{4}+\varepsilon, \frac{1}{4}+\varepsilon, 0, \frac{1}{4}-3\varepsilon), (1, 1, 2, 1, 3))$.

If $f^3(w',p') \in K_f$, then for each $k \in \{1,2,3\}$, $f^k(w',p')$ dominates $f^{k-1}(w',p')$ in expectation, and thus we have that $(w',p') \notin K_f$.

Next, we make the expectation f consistent such that $f^3(w',p') \in K_f$. Note that only a winning coalition in expectation who would be better off in the far-sighted core can move to other regions. Therefore, it suffices to construct the consistent expectation f such that some of players who change their region at $f^3(w',p')$ wind up being worse off at the final state in K_f.

For three distinct players $i,j,k \in \{1,2,3\}$, let $X(i;5) = \{(w,p) : w_j = w_k = \frac{1}{4} + \varepsilon,\ p_j = f_p^3(w',p')_j,\ p_k = f_p^3(w',p')_k,\ p_i = p_5 = 3,\ p_4 = 3 \text{ or } 1, \text{ and } \sum_{z \in p^3} w_z = \frac{2}{4} - 2\varepsilon\}$. Construct the expectation f such that every state in $X(i;5)$ is dominated in expectation by the state in D^* at which players j and k have halves. Then we have made the expectation f such that no state in $X(i;5)$ dominates $f^3(w',p')$ in expectation because any state that results from only player i changing her region winds up the state at which player i has zero.

For three distinct players $i,j,k \in \{1,2,3\}$, let $X(i,j;5) = \{(w,p) : w_k = \frac{1}{4} + \varepsilon,\ p_k = f_p^3(w',p')_k,\ p_i = p_j = p_5 = 3,\ p_4 = 3 \text{ or } 1, \text{ and } \sum_{z \in p^3} w_z = \frac{3}{4} - \varepsilon\}$. If $(w,p) \in X(1,2;5)$ satisfies that for some player z, $w_z > \frac{1}{2}(\frac{3}{4} - \varepsilon)$, then we make (w,p) dominated in expectation by the state $(\dot{w},p)$ with $\dot{w}_z = \frac{3}{4} - \varepsilon$ and $\dot{w}_3 = \frac{1}{4} + \varepsilon$. Then $(\dot{w},p)$ is dominated in expectation by some state $(\dot{w}',p') \in D^*$ with $\dot{w}'_z = 1$.

If $(w,p) \in X(1,2;5)$ satisfies that $\frac{1}{2}(\frac{3}{4}-\varepsilon) \geq \max\{w_z : z \neq 3\} > \frac{1}{4}-\varepsilon$, then we make (w,p) dominated in expectation by the state in D^* at which players 3 and $\min\{z : w_z > \frac{1}{4}-\varepsilon\}$ have halves. If $(w,p) \in X(1,2;5)$ satisfies that $\max\{w_z : z \neq 3\} \leq \frac{1}{4}-\varepsilon$, then we make (w,p) dominated in expectation by the state in D^* at which player 3 has a half and players z and y such that $z, y \neq 3$ and $w_z + w_y > \frac{1}{4}-\varepsilon$ have quarters. To embody the expectation f, choose players z and y such that $5z+y = \min\{5l+m : \text{for two distinct players } l, m \neq 3,\ w_l + w_m > \frac{1}{4}-\varepsilon\}$, which is the first by lexicographic ordering. Similarly, we construct f such that no state in $X(1,3;5)$ or $X(2,3;5)$ dominates $f^3(w',p')$ in expectation.

For three distinct players $i,j,k \in \{1,2,3\}$, let $X(i,j;k) = \{(w,p) : w_5 = \frac{1}{4}-3\varepsilon$, $p_5 = 3$, $p_i = p_j = p_k = f^3_p(w',p')_k$, $p_4 = 1$ or $f^3_p(w',p')_k$, and $\sum_{z \in p^{f^3_p(w',p')_k}} w_z = \frac{3}{4}+3\varepsilon\}$. If $(w,p) \in X(1,2;3)$ satisfies that for some player z, $w_z > \frac{1}{2}(\frac{3}{4}+3\varepsilon)$, then we make (w,p) dominated in expectation by the state $(\dot{w},\dot{p})$ with $\dot{w}_z = \frac{3}{4}+3\varepsilon$. Then the state $(\dot{w}',\dot{p}') \in D^*$ with $\dot{w}'_z = 1$ dominates $(\dot{w},\dot{p})$ in expectation. If $(w,p) \in X(1,2;3)$ satisfies that $\frac{1}{2}(\frac{3}{4}+3\varepsilon) \geq \max\{w_z : z \neq 5\} > \frac{1}{4}+3\varepsilon$, then we make (w,p) dominated in expectation by the state in D^* at which players 5 and $\min\{z : w_z > \frac{1}{4}+3\varepsilon\}$ have halves. If $(w,p) \in X(1,2;3)$ satisfies that $\#\{z : z \neq 5$ and $\frac{1}{4}+3\varepsilon \geq w_z > \frac{1}{4}(\frac{3}{4}+3\varepsilon)\} = 3$, then we make (w,p) dominated in expectation by the state $(\dot{w},\dot{p})$ such that players z and y who satisfy $5z+y = \min\{5l+m :$

for two distinct players $l, m \in \{1,3,4\}$, $w_l + w_m > \frac{1}{2}(\frac{3}{4} + 3\varepsilon)\}$ have $\frac{1}{2}(\frac{3}{4} + 3\varepsilon)s$ and make $(\dot{w}, \dot{p})$ dominated in expectation by the state in D^* such that players z and y have halves. In this case, one of players 1 and 2 ends up being worse off. If $(w,p) \in X(1,2;3)$ satisfies that $\max\{w_z : z \neq 5\} \leq \frac{1}{4} + 3\varepsilon$ and $\#\{z : w_z > \frac{1}{4}(\frac{3}{4} + 3\varepsilon)\} \leq 2$, then we make (w,p) dominated in expectation by the state $(\dot{w}, \dot{p})$ such that player i who has the least number among the wealthiest players at region 2, that is, $i = \min\{z \in p^2 :$ for any player $y \in p^2$, $w_z \geq w_y\}$, has $\frac{1}{2}(\frac{3}{4} + 3\varepsilon)$ and other two players z and y that satisfy $5z + y = \min\{5l + m : l, m \notin \{i, 5\}$, $l \neq m$, and $w_l, w_m \leq \frac{1}{4}(\frac{3}{4} + 3\varepsilon)\}$ have $\frac{1}{4}(\frac{3}{4} + 3\varepsilon)s$. Then we make $(\dot{w}, \dot{p})$ dominated in expectation by $(\dot{w}', \dot{p}')$ in D^* such that player i has a half and players z and y have quarters. Similarly, we construct f such that no state in $X(1,3;2)$ or $X(2,3;1)$ dominates $f^3(w', p')$ in expectation.

Finally, we construct the rest of the expectation f according to the way introduced in Theorem 40. Then we have that $K_f = D \cup \{f^3(w', p')\} \setminus \{(w', p')\}$, $(w', p') \in D$, and $f^3(w', p') \notin D$. Now, we only have to examine if f is consistent.

The expectation f is designed to satisfy that during the process of the expectation f, pillage can happen in each region at most once. If a player experiences pillage in his region, then he will never be pillaged during the rest of the process of f; that is, for any $(w,p) \in X \setminus K_f$ and any $i \in I$, if p_i satisfies that $\{z : f_w(w,p)_z \neq w_z\} \subset p^{p_i}$, then any positive integer k, $f_w^k(w,p)_i \geq f_w(w,p)_i$. It is easily seen that for any state $(w,p) \in X$ and any positive integer k, if $f^k(w,p) \neq f^{k-1}(w,p)$, then $i)$ there exists a region $r \in R$ such that $\{z : f^k_w(w,p)_z \neq f^{k-1}_w(w,p)_z\} \subset f^k_p(w,p)^r$; $ii)$ for all $q \neq r$, $f^k_p(w,p)^q = f^{k-1}_p(w,p)^q \setminus \{z : f^k_w(w,p)_z > f^{k-1}_w(w,p)_z\}$; and $iii)$ $\sum_{y\in\{z:f^k_w(w,p)_z > f^{k-1}_w(w,p)_z\}} f^{k-1}_w(w,p)_y > \sum_{y\in\{z:f^k_w(w,p)_z > f^{k-1}_w(w,p)_z\}} f^{k-1}_w(w,p)_y$. These show that $f(w,p)$ dominates (w,p) in expectation when $f(w,p) \neq (w,p)$ and (w,p) is undominated in expectation when $f(w,p) = (w,p)$. Therefore, f is consistent. □

2.4. Suggestion for further research

Throughout this paper, we have assumed that regions are connected with one another and thus players can travel from one region to another in one move. The results based on this assumption are meaningful in that they give general understanding of how spatial restriction affects stable distribution of wealth. Also, for applications, when we consider that many countries, which could be regarded as regions, are surrounded by the sea and we can travel from one country to another through the sea, the assumption seems to be an approximation to reality.

However, in order to describe real situations more exactly, we can generate a general model where some regions are not connected and thus players cannot travel between these regions in one move. A **geography correspondence** G embodies the general models as follows.

Definition 46. *A* ***geography correspondence*** *is a correspondence* $G : R \longrightarrow\!\!\longrightarrow R$ *satisfying for any* $r \in R$*, i)* $r \in G(r)$*; ii) if* $r' \in G(r)$ *then* $r \in G(r')$*; and iii) there exists a positive integer* k *such that* $G^k(r) = R$ *where* $G^k(r) = G^{k-1}(G(r))$*.*

For any $r \in R$, $G(r)$ denotes the regions that players at region r can go to in one move. Condition $i)$ means that players can stay in their regions. Condition $ii)$ means that connections between two regions are bilateral. And condition $iii)$ means that there is no separated region where players cannot travel. For example, we can define G as $G(1) = \{1, 2\}$, $G(2) = \{1, 2, 3\}$, and $G(3) = \{2, 3\}$, then G describes that three regions are located along a line.

The general model characterized by a geography correspondence is different from the previous model presented above in terms that weak players may be able to change states to defend themselves and thus some coalitions may not pillage less powerful coalitions. The following example shows how it works. Suppose that there are three players and five regions. Let G describe that five regions are located along a line, that is, $G(1) = \{1, 2\}, G(2) = \{1, 2, 3\}, ...,$ and $G(5) = \{4, 5\}$. Consider the state $(w, p) = ((\frac{4}{9}, \frac{1}{9}, \frac{4}{9}), (1, 3, 5))$, which expresses that players 1 and 3 have $\frac{4}{9}s$ while staying at region 1 and at region 5, respectively, and player 2 has $\frac{1}{9}$ while staying at region 3. At the state (w, p), player 2 can change the state to discourage another player who tried to pillage player 2. In cases that player

1 or players 1 and 3 approach player 2 to pillage, player 2 can move to region 4 to change the state so that player 3 can pillage player 2 alone. If player 3 alone pillages player 2, then player 3 has enough power to pillage player 1. Thus player 1 would not try to pillage player 2. In case that player 3 alone approaches player 2 to pillage, we can apply the same logic.

CHAPTER 3

Preference-based Cooperation in a Prisoner's Dilemma Game

The third chapter studies the possibility of cooperation based on players' preferences. Consider the following infinitely repeated game, similar to Ghosh and Ray (1996). At each stage, uncountable numbers of players are randomly matched without information about their partners' past actions and play a prisoner's dilemma game. The players have the option to continue their relationship and have the same discount factors. Also, they have two possible types: high ability player (H) or low ability player (L). H can produce better outcomes for its partner as well as for itself than L can. I look for an equilibrium that is robust against both pair-wise deviation and individual deviation and call such equilibrium a social equilibrium. I show that in this setting, long term cooperative behavior can happen in a social equilibrium. H wants to match and to play only with another H because an HH match produces better outcomes for H than an HL match. So H would break the match with an L to increase the possibility of meeting another H, and thus H would not play any cooperative action with an L. L knows this intention of H and realizes that L can only cooperate with another L. Consequently, both HH matches and LL matches are endowed with a scarcity value. This scarcity value is utilized by players to sustain cooperative relationships. Therefore, in a social equilibrium, players can play long term cooperative actions because of their preferences for their partners' types.

3.1. Introduction

This paper is a follow-up of Ghosh and Ray (1997). Ghosh and Ray showed that partial cooperative behavior in which some proportion of population cooperate with each other happens based on the structure of the model as well as based on a discount factor. Here, I model complete cooperative behavior in which whole population can cooperate with a particular kind of population based on the structure of the model as well as based on a discount factor.

It is well known in the economic literatures that a long term cooperative relationship can be sustained as an equilibrium in a variety of economic situations without any legal enforcement. Fudenberg and Maskin [4] show that a long term cooperative relationship is possible by an informal enforcement mechanism if the same players play the same substage game repeatedly. They say that if players have frequent and long term relationships, then the threat of future retaliation by the single opponent of a deviator can be strong enough to enforce cooperation.

Their argument relies on the assumption that players know their partner's history, because effective retaliation by the single opponent is based on knowing this history.

However, a lot of important long term relationships happen across a large population of players who are randomly matched and so have limited information on their partner's history. In such a situation, the future retaliation by the single opponent does not work well enough. Kandori [7] and Ellison [3] show that through contagious punishment, cooperative behavior is possible in random-matching games if there is some information flow across matches or there are some public means which expose information on other matches. But, their arguments still rely on the assumption that players could know a certain amount about their partner's history through the positive probability of re-matching with the old partners or public means.

Ghosh and Ray [6] find the assumption that there are some information flows across matches or public means which can be utilized by all players may not be valid in many economic contexts. They say "In some informal market in a developing country, a defaulter can move to another village or town where his past crime will almost certainly be unknown". Eventually, they propose the informal enforcement mechanism which can induce cooperative behavior as an equilibrium in the random-matching version of the game where players cannot know their partner's history at all.

This paper is a follow-up of Ghosh and Ray [6]. Their paper shows that if players have the option of continuing to play old opponents and the population is non-homogeneous in that some players are non-myopic with positive discount factor and others are myopic with zero discount factor, then cooperative behavior by the informal enforcement mechanism can happen as an equilibrium in intermediate economies where information flows and public means which expose information on other matches are absent. They explain that because the players can cooperate only with non-myopic type players, the existence of myopic type in the population endows the match with a non-myopic player with a scarcity value. This scarcity value is utilized to sustain cooperation among non-myopic players. That is, cooperative behavior relies on the assumption that a significant proportion of players in population are not concerned with their future payoff at all or just little. But, in a large class of interesting and important economic situations, most players in population are concerned with their own future payoff in some degree. This paper explores cooperation possibility in intermediate economies when all players are concerned with their future payoff.

We show that cooperative behavior by an informal enforcement mechanism is still possible in intermediate economies under fairly plausible conditions. First, we assume players are asymmetric in terms of their cooperation ability. So, there are two types of players in the model, high ability players and low ability players. It is well recognized that people are different with respect to their ability to improve an outcome in most economic situations. Some persons produce more and better goods with the same effort and the same time than others do. When it comes to cooperation, we might think that some coworkers have better cooperation ability, so we can achieve a better outcome when we work with them than when we work with others. We model the difference of these cooperation abilities. In this model, a player's payoff depends on his partner's type and his own type as well as his partner's cooperation level and his own cooperation level. But, all players share the same discount factor δ for their future payoff. This assumption distinguishes this paper from Ghosh and Ray [6]. In addition, we adopt some assumptions in Ghosh and Ray [6]. Second, Players still have the option of continuing to play old opponents. Third, players can propose only Match-wise Symmetric Norms which are defined to be a pure strategy profile in which players in the same kind of match use the same sequence of actions until one of players in the pair breaks the profile of pure strategies. We will call this match-wise symmetric norm a Social Norm. Besides the assumptions above, we restrict our interest to steady states defined later and the social norms which can induce the result similar to ones of Ghosh and Ray [6].

We follow the equilibrium concepts that are proposed in Ghosh and Ray [6]. So, the equilibrium in our model is robust against two kinds of deviations, individual deviation and pair-wise deviation which means joint deviation of players in the same match. The information in the model is not perfect because the players don't know what happens in other matches they are not involved in. Consequently, the whole game is its only subgame. For this reason, we adopt a Sequential Equilibrium in order to check if there is any player who prefers deviating individually to following a social norm off the equilibrium path as well as on the equilibrium path. In addition, to check if there is any pair where the players prefer pair-wise deviation from a social norm, we adopt the Negotiation-proofness concept which is reminiscent of "Bilateral rationality" in Ghosh and Ray [6]. That is, we try to find equilibrium such that it is proof against a negotiation. We formalize the negotiation process in the model which is not in Ghosh and Ray [6].

One of the main contributions of this paper to economic literature is constructing an example where cooperation between self-interested players can be expected in intermediate economies. In the model, the players can find out their partner's type with respect to cooperation ability. Because of a low expected payoff through cooperation or a high expected payoff through deviation, the players may not cooperate with a particular type of player(e.g. a high ability player or a low ability player). Therefore, a match with the other type of player(e.g. a low ability player or a high ability player) could be endowed with the scarcity value. This scarcity value makes the retaliation of breaking up the relationship fatal, where cooperation in the relationship is possible. That is, the players can utilize the effective retaliation of terminating relationship to sustain the cooperation with the type of player.

We need to explain how the low expected cooperation payoff or the high expected deviation payoff can block the cooperation between particular types of players. Suppose the expected payoff achieved through the cooperation between an H-player, which is short for a high ability player, and an L-player, which is short for a low ability player, is sufficiently small compared with the expected payoff achieved through the cooperation between two H-players. Thus, an H-player wants to make a long term cooperative relationship with an H-player. So, an H-player wants to break up the relationship with an L-player to increase the possibility of meeting an H-player. Because there is no information flow across matches and a random matching happens among a continuum of players, a player's past action does not affect his future expected payoff. According to the setting of the prisoner's dilemma game, an H-player who is matched with an L-player will not play any cooperative action. An L-player knows all this information because he knows the H-player's payoff system. Therefore, an L-player realizes that he himself can make a cooperative relationship only with an L-player. Consequently, the match between two L-players, as well as the match between two H-players, is endowed with the scarcity value. This can be one of the reasons why the cooperative relationship happens between players with the same cooperation ability.

The high expected deviation payoff can explain why the cooperative relationship happens between players with the different cooperation abilities. Suppose a single deviator in the match between two H-players can achieve such a high payoff in one stage that H-players cannot sustain the long term cooperative relationship between themselves. In addition, we suppose that the expected payoff achieved through the cooperation between an H-player and an L-player is so high that an L-player wants to make the long term cooperative relationship only with an H-player. According to the same reasoning as before, the match between an H-player and an L-player is endowed with the scarcity value which is utilized to sustain the long term cooperative relationship.

Another contribution of this paper is to suggest an efficient equilibrium. Most literatures that model the option of continuing old relationship assume that the players break up their relationship if their partner deviates from the prevailing profile of strategies. Such an assumption makes the expected payoff on the equilibrium path greatest. However, players may want to continue to cooperate even if one of the players in the pair deviates because the possibility of matching with another type of player at the next stage makes the present match valuable. In this paper, we suggest an equilibrium where, as long as the expected payoff on the equilibrium path would not be hurt, players continue their relationship although a single player in the pair deviated in previous stages.

There are three main results. First, we find the conditions under which a social equilibrium exists. Second, we prove that a social equilibrium must be unique if it exists. It is because of the negotiation-proofness which is one of our equilibrium concepts. The negotiation-proofness requires a unique possible cooperation pattern and the best cooperation levels in the model according to which players cooperate on the equilibrium path. Third, we show that even though players are sufficiently patient and their relationship is stable enough, players may not cooperate to the full cooperation level; that is, we can not apply a kind of Folk theorem to this model. This is because a deviator has the option to breaking up his relationship and can start a game with another partner. The worst expected punishment payoff of the deviator is limited to the expected payoff in the pool of unmatched players. In this model, the difference between the expected payoff of the matched players with cooperative relationship and the expected payoff of unmatched players relies on the proportion of players whom players can cooperate with in the pool of unmatched players. If the proportion of those players in the pool of unmatched players is high, then the difference between expected payoffs may be too small to punish the deviator severely. So, this limitation of the punishment may block players from fully cooperating with their partner although they are sufficiently patient and their relationship is stable enough.

Datta [2] studies a model similar to ours. He shows that cooperative behavior is possible by means of slowly building trust. However, equilibria in Datta [2] are not immune against pair-wise deviation as indicated by Ghosh and Ray [6]. In this context, people make a pair and play a repeated game between themselves. So, it is probable that they would have pre-play dialogue, or the negotiation before playing. Therefore, we think the equilibrium which is proof against the negotiation process, which is one of our equilibrium concepts, is realistic.

Watson [9] also examines the random-matching game of the prisoner's dilemma setting. In his model, player's type is characterized by their payoff structure similar to ours. But, a players' payoff in his model does not rely on their partner's type. This is different from our crucial assumption that players' payoff relies on their partner's type. The asymmetric player setting in Watson [9] is similar to the setting of different discount factors in Ghosh and Ray [6]. Different payoff settings cause players to have different degrees of preferences on cooperation as in Ghosh and Ray [6]. Watson [9] shows results similar to Datta [2]; that is, cooperation slowly develops.

The paper is organized as follows. In Section 2, we describe the model in detail. Here, we provide a time line. In Section 3, we present an assumption and equilibrium concepts. Also we mention the restriction on our interests. In section 4, we explore the general properties of a social equilibrium. In section 5, we provide the result about the social equilibrium. In section 6, we examine the feasibility of a steady state assumption and discuss some possible extensions of the model. The section 7 concludes.

3.2. The Model

3.2.1. Description of the game

There is a continuum of players in each group. There are two groups of players and each group has the same mass. Each player has his own type which is either high ability type or low ability type and player's type is determined by Nature regardless of player's preference. Players can not change their type. Players play an infinitely repeated game. At each stage every unmatched player is randomly matched only with an unmatched player in the different group. That is, random matching can happen only between unmatched players across groups. Matching makes a pair of two players. When players are matched, they have a chance to negotiate according to the negotiation process below before they start an extended version of the prisoner's dilemma game between themselves.

The negotiation process consists of at most five substages. But, there is no time discount during the negotiation process. At the first substage of the negotiation process, players in a pair can propose a negotiation to his partner. That is, players have two actions, "proposing a negotiation" or "not proposing a negotiation" and they in the pair choose these actions simultaneously. If both players in the pair propose a negotiation, then the player who is from group one makes an offer of a joint change from an original profile of strategies at the second substage. At the third substage, the player in the pair who is from group two chooses whether or not to consent to the offer. At the fourth and the fifth substages, they in the pair switch the turn and repeat the second and the third substages with proposing by the player from group two. If at least one of players in the pair does not propose a negotiation or they finish the negotiation process, then they start to play an extended version of the prisoner's dilemma game.

Each stage of the game consists of two substages at which players know all actions which have happened in their own matching games. At the first substage of each stage, each two players in the same pair plays an extended version of the prisoner's dilemma game with a continuum of actions $e \in [0, 1]$. At the second substage of the game, after watching the results at the first substage, players decide whether to continue their relationship to the next stage. Only if both players accede to continuing their relationship, can they play two substage game between themselves once more in the next stage; that is, if at least one of the players in the pair wants to terminate his relationship, then their relationship will be broken. But, even though both of them agree to continue their relationship, their relationship might be broken with the rate $\theta \in [0, 1]$ for some reasons outside of the model. In our daily life, we naturally think our relationship with anyone can be broken someday even though the relationship is profitable to all persons in the relationship. In addition, this exogenous relationship-breaking may happen independently of player's type.

There are many situations where players have to consider the possibility of breaking up their relationship against their will. First, a player may consider his partner's mistake. If he thinks his partner makes mistake with probability of $\theta \in [0, 1]$, then he can think his relationship will be broken with probability of θ when he expects his partner wants to keep their relationship. But, he does not need to consider partner's mistake when he and his partner choose a cooperation action because he can think mistakes in cooperation action are internalized in their cooperation action. If players have only two possible actions, Cooperation and Betrayal, and they use mixed strategy which can be recognized by their partner, then they can play more cooperation actions to compensate their mistakes. Therefore, he only needs to consider his partner's mistake about the termination/continuation decision. Consequently, an exogenous breaking possibility θ makes this model more realistic.

If their relationship is broken, then they return to the pool of unmatched players in each group and they will be equally likely matched with only one player in the pool of unmatched players of the different group at the beginning of the next stage. Therefore, if there are $\pi \in [0,1]$ proportion of high ability type players in the pool of unmatched players of the other group, then he will meet a high ability type player with the probability of π. My study focuses on a steady state where each π in both groups is constant. In addition, we assume that it is same across groups, and π is common knowledge.

Every player is either H-player which is short for high ability type player or L-player which is short for low ability type player. Player's type is determined by Nature regardless of player's preference. Player's type is common knowledge. Players' types in a pair determine the level of payoffs. The payoff functions are as follows; Here, e denotes his cooperation level and e' denotes his partner's cooperation level, when H-player meets H-player, the payoff of H-player is $r_{hh} : [0,1] \times [0,1] \rightarrow \mathbb{R}$ defined by $r_{hh}(e,e') = (1-e)e'c_{hh} + ee'a_{hh} + e(1-e')b_{hh}$, when H-player meets L-player, the payoff of H-player or L-player is $r_{hl} : [0,1] \times [0,1] \rightarrow \mathbb{R}$ defined by $r_{hl}(e,e') = (1-e)e'c_{hl} + ee'a_{hl} + e(1-e')b_{hl}$, when L-player meets L-player, the payoff of L-player is $r_{ll} : [0,1] \times [0,1] \to \mathbb{R}$ defined by $r_{ll}(e,e') = (1-e)e'c_{ll} + ee'a_{ll} + e(1-e')b_{ll}$, where $c_{hh} > a_{hh} > 0 > b_{hh}, c_{hl} > a_{hl} > 0 > b_{hl}, c_{ll} > a_{ll} > 0 > b_{ll}, a_{hh} > a_{hl} > a_{ll}$,and $c_{hh} \geq c_{hl} \geq c_{ll}$. So, in this model player can get payoffs according to his and his partner's type as well as according to their cooperation levels. By "$a_{hh} > a_{hl} > a_{ll}$" the model denotes high type players get more if they play the same level of co-operation and by "$c_{hh} \geq c_{hl} \geq c_{ll}$" the model denotes high ability players don't reduce the incentive for their partner to betray in the stage game compared with low ability players. These payoff functions are derived from the standard prisoner's dilemma game as

	cooperation	betrayal
cooperation	a_z, a_z	b_z, c_z
betrayal	c_z, b_z	$0, 0$

where $z \in \{hh, hl, ll\}$. Note that the payoffs of the cases that both players choose betrayal are normalized to '0'.

This is a complete information game because each player knows his partner's type and player's type is common knowledge. Player also knows his partner's action at the end of each substage. Of course, it is a perfect recall game. But, a player has no idea about what happened in other matches in which he is not involved. Finally, I assume that players are the same in terms of everything except their cooperation ability. So, every player has the same discount factor, $\delta_l = \delta_h = \delta$.

Note there is no discount among substages in a stage. The negotiation process and two substages of the extended version of the prisoner's dilemma game constitute the first stage of new match.

3.2.2. Time line

Here is the time line.

First, Nature chooses the players' group.

Second, Nature chooses the players' type.

Third, the players are randomly matched with a player in the other group.

Fourth, the players negotiate according to the negotiation process in each new match before they play the game.

Fifth, the players play an extended version of the prisoner's dilemma game.

At the first substage of the game, each chooses an action of cooperation level $e \in [0, 1]$. The payoffs are as follows;

$$\begin{aligned}
r_{hh}(e,e') &= (1-e)e'c_{hh} + ee'a_{hh} + e(1-e')b_{hh} \\
&\quad \text{when an H-player meets an H-player,} \\
r_{hl}(e,e') &= (1-e)e'c_{hl} + ee'a_{hl} + e(1-e')b_{hl} \\
&\quad \text{When an H-player meets an L-player or vice versa,} \\
r_{ll}(e,e') &= (1-e)e'c_{ll} + ee'a_{ll} + e(1-e')b_{ll} \\
&\quad \text{When an L-player meets an L-player,}
\end{aligned}$$

where e denotes a player's cooperation level and e' denotes his partner's cooperation level.

At the second substage of the game, the players choose a termination/continuation decision simultaneously. Only if both players in a pair choose the continuation decision, can they play the extended version of the prisoner's dilemma game between themselves at the next stage. But, even though both of them choose the continuation decision, their relationship might be broken with the rate $\theta \in [0, 1]$ for some exogenous reasons.

Sixth, the players whose relationship is broken up return to the pool of unmatched players and will be randomly matched with a player in the pool of unmatched players of the other group.

Note the future payoffs are discounted by $\delta \in (0, 1)$.

3.2.3. Personal History

The information set of each player can be described by the set of information which consists of; 1) actions in the negotiation process played by him and his partners, 2) his type and all his partners' types from the beginning on, 3) cooperation levels played by himself and his partners from the beginning on, 4) the termination/continuation decisions played by himself and his partners from the beginning on. I define the set of these four kinds of information as a personal history. Therefore, the personal history of a player is the records of his plays, his partners' plays, his type and his partners' types in all previous stages. A player can use only his personal history when he solves a stage game.

3.2.4. Strategy

The strategy of a player is a sequence of maps from his personal history to action in the negotiation process or to the action of a cooperation level and the termination/continuation decision subject to feasible condition. For example, "1) Play 'not Proposing a negotiation', 'Reject an offer', and 'Making an offer which contains $(e_{ij}, e_{ij}) = (1, 1)$ for the actions at the present match' in a negotiation process where $i, j \in \{h, l\}$, 2) Play '0' as cooperation action and Play 'T' (=break up the relationship) as the termination/continuation decision in the game" is a strategy.

3.2.5. Preference

All players are assumed to be maximizers of their own expected payoff.

3.3. Social Equilibrium

3.3.1. Assumption and Restrictions on Interest

Here, we restrict our interest in states and profiles of strategies. First, we are interested only in a steady state where the parameter π is constant over stages and a symmetric state where each group shares the same π. The feasibility of this restriction will be discussed at a later section.

Second, we only consider a "social norm", which means a profile of pure strategies in which players in the same kind of match use the same sequence of actions until one of the players in a pair deviates from the specified cooperation level in the profile of pure strategies. In our model, there are three kinds of matches; the match between H-player and H-player, the match between H-player and L-player, and the match between L-player and L-player. A social norm should indicate cooperation levels and the termination/continuation decisions for each kind of match. When we analyze the model, we need to know players' expected payoff in the pool of unmatched players with respect to a given profile of strategies because players choose their actions based on their expected payoff in the unmatched pool. If players could follow arbitrary profile of strategies, then we would have trouble to find players' expected payoff with respect to the given profile of strategies because we would not know shares between players in a pair. A social norm restriction excludes this trouble. In addition, a social norm can be reasonable because players in a pair will get the same payoff by playing a social norm on the equilibrium path.

Besides two restrictions on our interest, we need an assumption on actions off the equilibrium path and the offer in the negotiation process as mentioned before, because a social norm does not restrict actions off the equilibrium path and players can propose joint change which is not a social norm.

Assumption (1): Players propose only a social norm and in a punishment phase in which only one player in the pair deviates from the specified cooperation level in a social norm, the punisher does not enjoy worse per stage payoff than a deviator does, that is, $e_t \geq e'_t$ where e_t and e'_t are the deviator's action and the punisher's action respectively at the t^{th} stage in the punishment phase.

Off the equilibrium path, players can use alternate strategies that players in a pair play $(e = 1, e' = 0), (e = 0, e' = 1), (1, 0), \cdots$ because a social norm does not restrict actions off the equilibrium path. Both players in a pair can enjoy better expected payoff off the equilibrium path than on the equilibrium path. Assumption (1) excludes the case that players make use of punishment phase on purpose to maximize their payoffs.

3.3.2. Social Equilibrium

Now, we can propose two equilibrium concepts and define a "social equilibrium". First, the equilibrium should be a Sequential equilibrium.

Second, the equilibrium should satisfy negotiation-proofness which is defined as follows;

Negotiation-proofness: (I follow the definition of "Bilateral rationality" in Ghosh and Ray [6]) Let Σ be the set of all social norms. The social norm $\sigma^* \in \Sigma$ is said to satisfy "negotiation-proofness" if, given that all other players are following the norm, no matched pair of players can propose a joint change from the norm that increases their expected payoffs and satisfies the incentive constraint (which is defined as follows; the joint change from a social norm is said to satisfy the incentive constraint if, given that all other players will follow the social norm and one player in the pair follows the joint change from the social norm, the other player in the pair can not achieve a higher expected payoff by deviating from the joint change.)

Social Equilibrium: A social norm is said to be a "social equilibrium" if it is a sequential equilibrium which satisfies Negotiation-proofness.

3.4. General Properties of Social Equilibrium

Lemma 47. *If a player breaks up his present relationship, then his present action does not affect his expected payoff from the next stage on.*

Proof. Since there is no information flow across matches, the player's present action can have an effect on his expected payoff from the next stage on only when he meets his present partner, his present partner's new partners (whom his partner has met after his breaking up his relationship), or so on. Suppose the player meets his partner and breaks up his relationship at stage one. Let v_t be his expected value from stage two to stage $t \in \mathbb{N} \setminus \{1\}$ which is generated by playing with his present partner, his present partner's new partners, etc. from stage two to stage t. Let $\{v_t\}_{t=2}^{\infty}$ be the sequence of v_ts. Then, $\{v_t\}_{t=2}^{\infty}$ is bounded above by $\frac{\delta b}{1-\delta}$ and bounded below $\frac{\delta c}{1-\delta}$. The probability that he meets his partner and the players who have previously contacted with his partner, his partner's new partners etc. for finite stages is zero. Because there is a continuum of players in the pool of unmatched players and the players are randomly matched. By the limit location theorem in Mattuck [8], $\lim_{t\to\infty} \frac{\delta b}{1-\delta} p_t = \lim_{t\to\infty} \frac{\delta c}{1-\delta} p_t = 0$ since $\frac{\delta b}{1-\delta} p_t = \frac{\delta c}{1-\delta} p_t = 0$ for large t where p_t is the probability that he meets his partner or his partner's new partners, etc. for t stages. By the squeeze theorem for the limit of sequence in Mattuck [8], It is true that $\lim_{t\to\infty} v_t = 0$. That is, the player's present action does not affect his future expected payoff. □

Corollary 48. *In a social equilibrium, if a player breaks up his relationship on the equilibrium path, then both players in the match play cooperation level zero on the equilibrium path.*

Proof. By Lemma (1) and the setting of the prisoner's dilemma stage game, the result follows. □

Lemma 49. *If there exists a social equilibrium, then it should indicate constant cooperation level $e \in [0, 1]$ for the same kind of match (e.g. the match between two H-players) on the equilibrium path.*

Proof. Suppose there exists a social equilibrium in which players play according to the sequence $\{e_t\}_{t=1}^{\infty}$ for the specific kind of match on the equilibrium path where e_t denotes the action of a cooperation level at the t^{th} stage after the players are matched. By way of contradiction, suppose there exists $s \in \mathbb{N}$ such that $e_s \neq e_{s+1}$. Since "$e_s > 0$ or $e_{s+1} > 0$" and the sequence $\{e_t\}_{t=1}^{\infty}$ is on the equilibrium path, players in this kind of match will choose to continue their relationship on the equilibrium path by Corollary (2). Let $v(\{e_t\}_{t=s}^{\infty})$ be the expected payoff of the player who are in the same kind of match given that he and his partner in the match will play according to $\{e_t\}_{t=s}^{\infty}$ from the present stage on the equilibrium path and play according to the social equilibrium in other matches. If $v(\{e_t\}_{t=s}^{\infty}) > v(\{e_t\}_{t=s+1}^{\infty})$, then the joint change, $\{e_1, \cdots, e_s, e_s, e_{s+1}, \cdots\}$ from $\{e_1, \cdots, e_s, e_{s+1}, \cdots\}$, does not change the deviator's expected payoff and it increases the expected payoff of the follower of this joint change from the social equilibrium. Therefore, the joint change, $\{e_1, \cdots, e_s, e_s, e_{s+1}, \cdots\}$ satisfies the incentive constraint and increases both players' expected payoff. Similarly, if $v(\{e_t\}_{t=s}^{\infty}) < v(\{e_t\}_{t=s+1}^{\infty})$, then the joint change, $\{e_1, \cdots, e_{s-1}, e_{s+1}, \cdots\}$ from $\{e_1, \cdots, e_{s-1}, e_s, e_{s+1}, \cdots\}$, satisfies the incentive constraint and increases both players' expected payoff. If $v(\{e_t\}_{t=s}^{\infty}) = v(\{e_t\}_{t=s+1}^{\infty})$, then either $v(\{e_t\}_{t=s+1}^{\infty}) > v(\{e_t\}_{t=s+2}^{\infty})$ or $v(\{e_t\}_{t=s+1}^{\infty}) < v(\{e_t\}_{t=s+2}^{\infty})$ since $e_s \neq e_{s+1}$. Therefore, in this case we can see the same result from the conclusions above. Consequently, these contradict the definition of the social equilibrium. $\square$

Remark 50. *Suppose there exists a social equilibrium with* (e_{ij}) *for* $i, j \in \{h, l\}$ *as the profile of cooperation levels on the equilibrium path where* e_{ij} *denotes a constant cooperation level when an i-type player meets a j-type player. Let the players play according to the sequence* $\{k_t\}_{t=1}^{\infty}$ *for the specific kind of match on the equilibrium path where* k_t *denotes the action of the termination/continuation decision at the* t^{th} *stage after the players are matched. If there exists* $s \in \mathbb{N}$ *such that* $k_s = T$ *(= terminating the relationship), then* $e_{ij} = 0$ *for the match on the equilibrium path according to Corollary (2). If the players terminate their relationship during the match, the actions in the match do not affect the players' future expected payoff by Lemma (1). Therefore, players can terminate their relationship from the first stage of this match without any reduction of their expected payoff. That is, there exists a social equilibrium with the same* (e_{ij}) *such that players play according to the sequence* $\{k_t\}_{t=1}^{\infty} = \{T\}_{t=1}^{\infty}$ *for the match on the equilibrium path.*

According to Lemma (3) and Remark (4), in order to find out cooperation levels in a social equilibrium, we only have to consider the cases that players reach the maximum cooperation level at once and use the fixed the termination/continuation decisions for each match.

Notations: Let $e_{hh}, e_{hl} = e_{lh}, e_{ll} \in [0, 1]$ be the cooperation levels when H-player meets H-player, when H-player meets L-player or L-player meets

H-player, and when L-player meets L-player respectively. Let $k_{hh}, k_{hl} = k_{lh}, k_{ll} \in \{T, C\}$ be the termination/continuation decisions when H-player meets H-player, when H-player meets L-player or when L-player meets H-player, and when L-player meets L-player respectively where T denotes Terminating a relationship and C denotes continuing a relationship. For a notational convenience, let $E = (e_{hh}, e_{hl}, e_{ll}), K = (k_{hh}, k_{hl}, k_{ll}), E_h = (e_{hh}, e_{hl}), K_h = (k_{hh}, k_{hl}), E_l = (e_{hl}, e_{ll})$, and $K_l = (k_{hl}, k_{ll})$, then $E \in [0,1]^3, K \in \{T, C\}^3, E_i \in [0,1]^2, K_i \in \{T, C\}^2$ for $i \in \{h, l\}$. Define $V_{ij}(E_i, K_i) : [0,1]^2 \times \{T, C\}^2 \rightarrow \mathbb{R}$ to be a matched i-type player's expected payoff function when i-type player meets j-type player and he expects that from the present stage on he and all his partners play e_{hh} and e_{hl} or e_{ll} and e_{hl} as cooperation levels and k_{hh} and k_{hl} or k_{ll} and k_{hl} as the termination/continuation decisions where $i, j \in \{h, l\}$. Define $V_i(E_i, K_i) : [0,1]^2 \times \{T, C\}^2 \rightarrow \mathbb{R}$ to be an unmatched player's expected payoff function of i-type player and he expects that from the present stage on he and all his partners play E_i as cooperation levels and K_i as the termination/continuation decisions where $i \in \{h, l\}$. Then, we can derive that $V_i(E_i, K_i) = \pi_i V_{ii}(E_i, K_i) + (1 - \pi_i) V_{ij}(E_i, K_i)$ where π_i denotes the proportion of i-type players in the pool of unmatched players of the other group. Note we assume a symmetric state where each group shares same π and θ. Finally, define $V_{ij}^{b(1)}(E_i, K_i)$ and $V_{ji}^{p(1)}(E_j, K_j)$ to be a i-type deviator's expected payoff function in the first stage of the punishment phase, $V_{ij}^{d(1)} : [0,1]^2 \times \{T, C\}^2 \rightarrow \mathbb{R}^+$, and a j-type punisher's expected payoff function in the first stage of the punishment phase, $V_{ji}^{p(1)} : [0,1]^2 \times \{T, C\}^2 \rightarrow \mathbb{R}^+$, respectively when i-type player betrays j-type player and they expect that from the present stage on their all new partners and they in the new match play e_{hh} and e_{hl} or e_{ll} and e_{hl} as cooperation levels and k_{hh} and k_{hl} or k_{ll} and k_{hl} as the termination/continuation decisions on the equilibrium path where $i, j \in \{h, l\}$

Lemma 51. *Suppose there exists a social norm which is a sequential equilibrium with cooperation levels and the termination/continuation decisions,* $E = (e_{hh}, e_{hl}, e_{ll})$ *and* $K = (k_{hh}, k_{hl}, k_{ll})$*, on the equilibrium path. In the social norm, each player can cooperate with at most one type of player, that is, if* $e_{hl} > 0$ *then* $e_{hh} = e_{ll} = 0$*, if* $e_{hh} > 0$ *then* $e_{hl} = 0$*, and if* $e_{ll} > 0$ *then* $e_{hl} = 0$*.*

Proof. Suppose there exists a social norm which is a sequential equilibrium and it contains the profiles of actions $E = (e_{hh}, e_{hl}, e_{ll})$ and $K = (k_{hh}, k_{hl}, k_{ll})$ on the equilibrium path. Let $\{i, j\} = \{h, l\}$ (which means $i, j \in \{h, l\}$ and $i \neq j$). Suppose that $V_{ii}(E_i, K_i) \geq V_{ij}(E_i, K_i)$, then $V_{ii}(E_i, K_i) \geq V_i(E_i, K_i) \geq V_{ij}(E_i, K_i)$ by the definition of $V_i(E_i, K_i)$. Since it is a sequential equilibrium, it satisfies the incentive constraint below;

$$
\begin{aligned}
& r_{ij}(0, e_{ij}) + \max\{\delta(1-\theta)V_{ij}^{d(1)}(E_i, K_i) + \delta\theta V_i(E_i, K_i), V_i(E_i, K_i)\} \\
\leq \;& r_{ij}(e_{ij}, e_{ij}) + \max\{\delta(1-\theta)V_{ij}(E_i, K_i) + \delta\theta V_i(E_i, K_i), V_i(E_i, K_i)\}
\end{aligned}
$$

$$
\Longrightarrow r_{ij}(0, e_{ij}) + V_i(E_i, K_i) \leq r_{ij}(e_{ij}, e_{ij}) + V_i(E_i, K_i)
$$

Since $\max\{\delta(1-\theta)V_i^{d(1)}(E_i, K_i) + \delta\theta V_i(E_i, K_i), V_i(E_i, K_i)\} \geq V_i(E_i, K_i) \geq V_{ij}(E_i, K_i)$. Therefore, the only possible cooperation level e_{ij} is zero. Similarly, we can prove the case $V_{ij}(E_i, K_i) \geq V_{ii}(E_i, K_i)$. Consequently, this result follows. □

Remark 52. *If there exists a social equilibrium, then the profile of its cooperation levels is one of the solutions to the problems below. If there exists a social equilibrium, then in the social equilibrium each player plays the constant cooperation level for each match by Lemma (3). Without loss of generality we can assume that players play the fixed the termination/continuation decision for each match for the social equilibrium by Remark (4). The optimization problems below find out the maximum cooperation levels among cooperation levels of sequential equilibria when players play the constant cooperation level and the fixed the termination/continuation decision for each match. Therefore, if there exists a social equilibrium, then the profile of its cooperation levels should be one of the solutions* *to the problems below.* ***Optimization problem 1); When H-players cooperate only with H-players and L-players cooperate only with L-players,***

1)-1 H-players solve this;

$$\max_{e_{hh\in[0,1]}} V_{hh}(E_h, K_h) \equiv \max_{e_{hh\in[0,1]}} \frac{r_{hh}(e_{hh}, e_{hh}) + \delta\theta V_h(E_h, K_h)}{1-\delta(1-\theta)}$$

$$s.t.1) r_{hh}(0, e_{hh}) + \max\{\delta V_h(E_h, K_h), \delta(1-\theta)V_{hh}^{d(1)}(E_h, K_h) + \delta\theta V_h(E_h, K_h)\}$$

$$\leq \quad \frac{r_{hh}(e_{hh}, e_{hh}) + \delta\theta V_h(E_h, K_h)}{1-\delta(1-\theta)}$$

such that $V_{hh}^{d(1)}(E_h, K_h)$ *and* $V_{hh}^{p(1)}(E_h, K_h)$ *satisfy the incentive constraint and the participation constraint which are explained below where* $E_h = (e_{hh}, e_{hl} = 0)$ *and* $K_h = (k_{hh} = C, k_{hl} = T)$.

1)-2 L-players solve this;

$$\max_{e_{ll}\in[0,1]} V_{ll}(E_l, K_l) \equiv \max_{e_{ll}\in[0,1]} \frac{r_{ll}(e_{ll}, e_{ll}) + \delta\theta V_l(E_l, K_l)}{1-\delta(1-\theta)}$$

$$s.t. r_{ll}(0, e_{ll}) + \max\{\delta V_l(E_l, K_l), \delta(1-\theta)V_{ll}^{d(1)}(E_l, K_l) + \delta\theta V_l(E_l, K_l)\}$$

$$\leq \quad \frac{r_{ll}(e_{ll}, e_{ll}) + \delta\theta V_l(E_l, K_l)}{1-\delta(1-\theta)}$$

such that $V_{ll}^{d(1)}(E_l, K_l)$ *and* $V_{ll}^{p(1)}(E_l, K_l)$ *satisfy the incentive constraint and the participation constraint where* $E_l = (e_{hl} = 0, e_{ll})$ *and* $K_l = (k_{hl} = T, k_{ll} = C)$.

1)-3 Consistency condition;

$$V_h(E_h, K_h) = \pi V_{hh}(E_h, K_h) + (1-\pi)\delta V_h(E_h, K_h)$$

$$V_l(E_l, K_l) = \pi \delta V_l(E_l, K_l) + (1-\pi) V_{ll}(E_l, K_l).$$

Optimization problem 2); When H-players cooperate only with L-players and vice versa on the equilibrium path,

2)-1 all players solve this;

$$\begin{aligned}
\max_{e_{hl}\in[0,1]} V_{hl}(E_h, K_h) &\equiv \max_{e_{hl}\in[0,1]} \frac{r_{hl}(e_{hl}, e_{hl}) + \delta\theta V_h(E_h, K_h)}{1-\delta(1-\theta)} \\
\max_{e_{hl}\in[0,1]} V_{lh}(E_l, K_l) &\equiv \max_{e_{hl}\in[0,1]} \frac{r_{lh}(e_{hl}, e_{hl}) + \delta\theta V_l(E_l, K_l)}{1-\delta(1-\theta)}
\end{aligned}$$

$$\begin{aligned}
& s.t. 1)\ r_{hl}(0, e_{hl}) + \max\{\delta V_h(E_h, K_h), \delta(1-\theta)V_{hl}^{d(1)}(E_h, K_h) + \delta\theta V_h(E_h, K_h)\} \\
\leq\ & \frac{r_{hl}(e_{hl}, e_{hl}) + \delta\theta V_h(E_h, K_h)}{1-\delta(1-\theta)} \\
& 2)\ r_{lh}(0, e_{hl}) + \max\{\delta V_l(E_l, K_l), \delta(1-\theta)V_{lh}^{d(1)}(E_l, K_l) + \delta\theta V_l(E_l, K_l)\} \\
\leq\ & \frac{r_{lh}(e_{lh}, e_{lh}) + \delta\theta V_l(E_l, K_l)}{1-\delta(1-\theta)}
\end{aligned}$$

such that $V_{ij}^{d(1)}(E_i, K_i)$ *and* $V_{ji}^{p(1)}(E_j, K_j)$ *satisfy the incentive constraint and the participation constraint where* $E_h = (e_{hh} = 0, e_{hl}), K_h = (k_{hh} = T, k_{hl} = C), E_l = (e_{hl}, e_{ll} = 0),\ K_h = (k_{hl} = C, k_{ll} = T),$ *and* $\{i, j\} = \{h, l\}$.

2)-2 Consistency condition;

$$V_h(E_h, K_h) = \pi\delta V_h(E_h, K_h) + (1-\pi)V_{hl}(E_h, K_h)$$

$$V_l(E_l, K_l) = (1-\pi)V_{lh}(E_l, K_l) + \pi\delta V_l(E_l, K_l).$$

Note two points; First, the solutions to the problems above satisfy necessary conditions to become a social equilibrium, but the solutions may not satisfy sufficient conditions. When players solve the original optimization problem, they regard the expected payoffs, $V_{ij}(E_i, K_i)$, in the problem as constants and try to find the fixed points of $V_{ij}(E_i, K_i)$. If there exists a social equilibrium, then the social equilibrium should be the maximum cooperation levels among cooperation levels of sequential equilibria. So, the social equilibrium should be the solution of the problems above. But, in the problems above the expected payoffs, $V_{ij}(E_i, K_i)$, are functions of cooperation levels. Therefore, the solutions to the problems above may not satisfy negotiation-proofness. Second, constraints in each problem denote the Incentive constraint. The incentive constraint means if all his partners follow the social norm, then he is better off when he also follows the social norm. We don't need to state because it is the part of consistency condition. The participation constraint means players don't want to terminate his present relationship in the social norm.

Lemma 53. *Suppose there exists a social equilibrium with a deviator's expected payoff function in the first stage of the punishment phase,* $V_{ij}^{d(1)}(E_i^*, K_i^*)$ *for* $i, j \in \{h, l\}$*. If* $E^* = (e_{hh}, e_{hl}, e_{ll})$ *is the profile of cooperation levels on the equilibrium path with the deviator's expected payoff function in the first stage of the* *punishment phase, $V_{ij}^{d(1)}(E_i^*, K_i^*)$. Then, the same profile E^* for the equilibrium path can be derived with the deviator's expected payoff function in the first stage of the punishment phase, $V_i(E_i^*, K_i^*)$ for $i, j \in \{h, l\}$.*

Proof. If E^* is derived with $V_{ij}^{d(1)}(E_i^*, K_i^*) \leq V_i(E_i^*, K_i^*)$ for $i, j \in \{h, l\}$, then the change of the deviator's expected payoff function in the first stage of the punishment phase to $V_i(E_i^*, K_i^*)$ does not make any difference to the players in terms of cooperation levels. In case $V_{ij}^{d(1)}(E_i^*, K_i^*) < V_i(E_i^*, K_i^*)$, the deviator will terminate his relationship after deviating. Although players change the deviator's expected payoff function from $V_{ij}^{d(1)}(E_i^*, K_i^*)$ to $V_i(E_i^*, K_i^*)$, they are faced with the same optimization problems in Remark (3). So, we only need to show that we can change the deviator's expected payoff function in the first stage of the punishment phase in case $V_{ij}^{d(1)}(E_i^*, K_i^*) > V_i(E_i^*, K_i^*)$ without changing the equilibrium cooperation levels E^*. By way of contradiction, we suppose not. Then, there exists a social equilibrium with $V_{ij}^{d(1)}(E_i^*, K_i^*) > V_i(E_i^*, K_i^*)$ in which cooperation levels would change if we would change $V_{ij}^{d(1)}(E_i^*, K_i^*)$ to $V_i(E_i^*, K_i^*)$. Let $e_{ij}(= e_{ji}) > 0$. Note that Remark (3) denotes the necessary conditions to become a social equilibrium. From constraints in Remark (3), $(E^*, K^*) = ((e_{hh}, e_{hl}, e_{ll}), (k_{hh}, k_{hl}, k_{ll}))$ satisfies the following;

$$
\begin{aligned}
& 1)\ r_{ij}(0,e_{ji}) + \max\{\delta V_i(E_i,K_i), \delta(1-\theta)V_{ij}^{d(1)}(E_i^*,K_i^*) + \delta\theta V_i(E_i^*,K_i^*)\} \\
\leq\ & \frac{r_{ij}(e_{ij},e_{ji}) + \delta\theta V_i(E_i^*,K_i^*)}{1-\delta(1-\theta)} \\
& 2)\ r_{ji}(0,e_{ij}) + \max\{\delta V_j(E_j,K_j), \delta(1-\theta)V_{ji}^{d(1)}(E_j^*,K_j^*) + \delta\theta V_j(E_j^*,K_j^*)\} \\
\leq\ & \frac{r_{ji}(e_{ji},e_{ij}) + \delta\theta V_j(E_j^*,K_j^*)}{1-\delta(1-\theta)}.
\end{aligned}
$$

If we change $V_{ij}^{d(1)}(E_i^*,K_i^*)$ to $V_i(E_i^*,K_i^*)$, then the constraints become

$$
1)\ r_{ij}(0,e_{ji}) + \delta V_i(E_i^*,K_i^*) \leq \frac{r_{ij}(e_{ij},e_{ji}) + \delta\theta V_i(E_i^*,K_i^*)}{1-\delta(1-\theta)} \tag{1}
$$

$$
2)\ r_{ji}(0,e_{ij}) + \delta V_j(E_j^*,K_j^*) \leq \frac{r_{ji}(e_{ji},e_{ij}) + \delta\theta V_j(E_j^*,K_j^*)}{1-\delta(1-\theta)}. \tag{2}
$$

(e_{ij},e_{ji}) still satisfy the constraints (1) and (2) above since $V_{ij}^{d(1)}(E_i^*,K_i^*) > V_i(E_i^*,K_i^*)$. The objective functions in Remark (3) require as high cooperation levels as possible such that the cooperation levels satisfy constraints of the problem. This means players could derive greater cooperation levels (e'_{ij},e'_{ji}) than (e_{ij},e_{ji}) if they would change their deviator's expected payoff function in the first stage of the punishment phase from $V_{ij}^{d(1)}(E_i^*,K_i^*)$ to $V_i(E_i^*,K_i^*)$. Since $V_i(E_i^*,K_i^*)$ is a function $V_i : [0,1]^2 \times \{T,C\}^2 \to [b_{hl},c_{hl}] \subset \mathbb{R}$ and the players can choose it as the deviator's expected payoff function in the first stage of the punishment phase, the social norm containing E^* does not satisfy negotiation-proofness of the social equilibrium. This contradiction results from the assumption that there exists a social equilibrium with $V_{ij}^{d(1)}(E_i^*, K_i^*) > V_i(E_i^*, K_i^*)$ in which cooperation levels would change if we would change $V_{ij}^{d(1)}(E_i^*, K_i^*)$ to $V_i(E_i^*, K_i^*)$. □

3.5. Results

Through these Lemmas, Remarks, and Claim above, we can derive the following results.

Proposition (1): There exists a social equilibrium if and only if the profile of cooperation levels $E^* = (e_{hh}^*, e_{hl}^*, e_{ll}^*)$ derived in the Remark (6) with $V_{ij}^{d(1)}(E_i^*, K_i^*) = V_i(E_i^*, K_i^*)$ satisfies two conditions below; Let $k^* = (k_{hh}^*, k_{hl}^*, k_{ll}^*)$ be the profile of the termination/continuation decisions associated with E^* in Remark (6). Condition 1), if $0 \leq e_{hl}^* < 1$, then E^* satisfies either

$$\begin{aligned} r_{hl}(0,e) + \delta V_h(E_h^*, K_h^*) &> \frac{r_{hl}(e,e) + \delta\theta V_h(E_h^*, K_h^*)}{1-\delta(1-\theta)} \\ \textit{or } r_{hl}(0,e) + \delta V_l(E_l^*, K_l^*) &> \frac{r_{hl}(e,e) + \delta\theta V_l(E_l^*, K_l^*)}{1-\delta(1-\theta)} \end{aligned}$$

or both for all $e \in (e_{hl}^*, 1]$. Condition 2), if $0 \leq e_{ii}^* < 1$ for $i \in \{h, l\}$, then E^* satisfies

$$r_{ii}(0, e) + \delta V_i(E_i^*, K_i^*) > \frac{r_{ii}(e, e) + \delta\theta V_i(E_i^*, K_i^*)}{1 - \delta(1 - \theta)}$$

for all $e \in (e_{ii}^*, 1]$.

Proof. (Necessary condition) Suppose that there exists a social equilibrium. The profile of cooperation levels derived in Remark (6) with $V_{ij}^{d(1)}(E_i^*, K_i^*) = V_i(E_i^*, K_i^*)$ is the maximum cooperation levels which are possible in the sequential equilibrium with the constant cooperation level and the fixed the termination/continuation decision for each match. So, if there exists a social equilibrium, then it should be derived as in the Remark (6). Conditions 1) and 2) above denote the negotiation-proofness. Since a social equilibrium satisfies negotiation-proofness, a social equilibrium should satisfy conditions 1) and 2) above.

(Sufficient condition) Suppose $E = (e_{hh}, e_{hl}, e_{ll})$ derived in the Remark (6) with $V_{ij}^{d(1)}(E_i^*, K_i^*) = V_i(E_i^*, K_i^*)$ satisfies the two conditions above. The profile of actions E is the part of a sequential equilibrium by the constraints in Remark (6). Therefore in order to prove the sufficient condition, we only need to show that players can not increase their cooperation levels from the profile of actions E. According to conditions 1) and 2), any cooperation level which is greater than the cooperation level in E breaks the incentive constraint. Consequently, we can construct the social equilibrium which contains E. $\square$

Remark 54. *We can express conditions 1) and 2) in Proposition (1) as follows;*

Condition 1'), if $0 = e^*_{hl}$*, then* E^* *satisfies either*

$$\begin{aligned} ec_{hl} - \frac{r_{hl}(e,e)}{1-\delta(1-\theta)} &> \frac{-\delta\pi(1-\theta)r_{hh}(e^*_{hh},e^*_{hh})}{(1-\delta(1-\theta))(1-\delta+\delta\theta+\delta\pi-\delta\theta\pi)} \\ or \quad ec_{hl} - \frac{r_{hl}(e,e)}{1-\delta(1-\theta)} &> \frac{-\delta(1-\pi)(1-\theta)r_{ll}(e^*_{ll},e^*_{ll})}{(1-\delta(1-\theta))(1-\delta\pi+\delta\theta\pi)} \end{aligned}$$

or both for all $e \in (0,1]$ *and if* $0 < e^*_{hl} < 1$*, then* E^* *satisfies*

$$\frac{\partial r_{hl}(0,e)}{\partial e}\,|_{e=e^*_{hl}} \geq \frac{\partial r_{hl}(e,e)}{(1-\delta(1-\theta))\partial e}\,|_{e=e^*_{hl}}$$

$\Longleftrightarrow$

$$\begin{aligned} \delta(1-\theta)(1-2\pi)c_{hl} &\geq -b_{hl} \\ or \quad \delta(1-\theta)(2\pi-1)c_{hl} &\geq -b_{hl}. \end{aligned}$$

Condition 2'), if $0 = e^*_{ii}$ *for* $i \in \{h,l\}$*, then* E^* *satisfies*

$$ec_{ii} - \frac{r_{ii}(e,e)}{1-\delta(1-\theta)} > \frac{-\delta(1-\pi_i)(1-\theta)r_{hl}(e^*_{hl},e^*_{hl})}{(1-\delta(1-\theta))(1-\delta\pi_i+\delta\theta\pi_i)}$$

for all $e \in (0,1]$ *and if* $0 < e^*_{ii} < 1$ *for* $i \in \{h,l\}$, *then* E^* *satisfies*

$$\frac{\partial r_{ii}(0,e)}{\partial e} \mid_{e=e^*_{ii}} \geq \frac{\partial r_{ii}(e,e)}{(1-\delta(1-\theta))\partial e} \mid_{e=e^*_{ii}}$$

$\iff$

$$\delta(1-\theta)(1-2\pi_i)c_{ii} \geq -b_{ii}$$

where $\pi_h = \pi$ *and* $\pi_l = (1-\pi)$.

Remark 55. *According to Proposition (1), if there exists a social equilibrium, then we can figure out the cooperation levels as follows;*

When i-type players cooperates only with i-type players *where* $i = H, L$.

$$\text{(1)} \qquad V_h(E_h,K_h) = \pi V_{hh}(E_h,K_h) + (1-\pi)\delta V_h(E_h,K_h) = \frac{\pi V_{hh}(E_h,K_h)}{1-\delta(1-\pi)},$$

from the definition of the function $V_{hh}(E_h,K_h)$,

$$V_{hh}(E_h,K_h) \equiv \frac{r_{hh}(e_{hh},e_{hh}) + \delta\theta V_h(E_h,K_h)}{1-\delta(1-\theta)} = \frac{r_{hh}(e_{hh},e_{hh}) + \delta\theta\frac{\pi V_{hh}(E_h,K_h)}{1-\delta(1-\pi)}}{1-\delta(1-\theta)}$$

$$\iff \quad V_{hh}(E_h, K_h)\frac{(1-\delta+\delta\theta)(1-\delta+\delta\pi)-\delta\theta\pi}{1-\delta(1-\pi)}$$
$$= \quad (1-e_{hh})e_{hh}c_{hh} + (e_{hh})^2 a_{hh} + e_{hh}(1-e_{hh})b_{hh}$$

$$\iff \quad V_{hh}(E_h, K_h) \qquad \textit{(2)}$$
$$= \quad \frac{(1-\delta(1-\pi))((1-e_{hh})e_{hh}c_{hh} + (e_{hh})^2 a_{hh} + e_{hh}(1-e_{hh})b_{hh})}{(1-\delta+\delta\theta)(1-\delta+\delta\pi)-\delta\theta\pi}$$
$$= \quad \frac{(1-\delta(1-\pi))r_{hh}(e_{hh}, e_{hh})}{(1-\delta)(1-\delta+\delta\theta+\delta\pi-\delta\theta\pi)}.$$

From constrains in the problem 1) of the Remark (6) and equations (1) and (2) above,

$$c_{hh}e_{hh} + \frac{\delta\pi V_{hh}(E_h, K_h)}{(1-\delta(1-\pi))} \leq V_{hh}(E_h, K_h)$$
$$\iff V_{hh}(E_h, K_h)\frac{\delta(1-\theta)(\delta-1)(1-\pi)}{1-\delta(1-\pi)} \leq (e_{hh})^2(a_{hh} - c_{hh} + b_{hh}) + e_{hh}b_{hh}$$

$$\iff \quad V_{hh}(E_h, K_h) = \frac{(1-\delta(1-\pi))r_{hh}(e_{hh}, e_{hh})}{(1-\delta)(1-\delta+\delta\theta+\delta\pi-\delta\theta\pi)}$$
$$\geq \quad \frac{(1-\delta(1-\pi))(r_{hh}(e_{hh}, e_{hh}) - e_{hh}c_{hh})}{\delta(1-\theta)(\delta-1)(1-\pi)}$$

$$\iff \quad \frac{r_{hh}(e_{hh}, e_{hh})}{(1-\delta+\delta\theta+\delta\pi-\delta\theta\pi)} \geq -\frac{r_{hh}(e_{hh}, e_{hh}) - e_{hh}c_{hh}}{\delta(1-\theta)(1-\pi)}$$
$$\iff \quad e_{hh}\{e_{hh}(a_{hh} - c_{hh} - b_{hh}) + b_{hh} + \delta(1-\theta)(1-\pi)c_{hh}\} \geq 0 \ \textit{and} \ e_{ll} \in [0, 1].$$

Then, we can find out the solution $e_{hh,1} \in [0,1]$ *to this problem as follows;* $e_{hh,1} =$

$$= 1 \text{ if } 0 < a_{hh} - c_{hh} - b_{hh}, -\left(\frac{b_{hh} + \delta(1-\theta)(1-\pi)c_{hh}}{a_{hh} - c_{hh} - b_{hh}}\right) \leq 1 \quad \textit{(hh-1)}$$

$$= 0 \text{ if } 0 < a_{hh} - c_{hh} - b_{hh}, 1 < -\left(\frac{b_{hh} + \delta(1-\theta)(1-\pi)c_{hh}}{a_{hh} - c_{hh} - b_{hh}}\right) \quad \textit{(hh-2)}$$

$$= 1 \text{ if } 0 = a_{hh} - c_{hh} - b_{hh}, 0 \leq b_{hh} + \delta(1-\theta)(1-\pi)c_{hh} \quad \textit{(hh-3)}$$

$$= 0 \text{ if } 0 = a_{hh} - c_{hh} - b_{hh}, b_{hh} + \delta(1-\theta)(1-\pi)c_{hh} < 0 \quad \textit{(hh-4)}$$

$$= \min\{\max\{0, -\left(\frac{b_{hh} + \delta(1-\theta)(1-\pi)c_{hh}}{a_{hh} - c_{hh} - b_{hh}}\right)\}, 1\} \text{ if } a_{hh} - c_{hh} - b_{hh} < 0 \quad \textit{(hh-5)}$$

Similarly, we can derive the solution $e_{ll,1}$ *to the problem 1) of the Remark (6) for L-players as follows;*

$$e_{ll}\{e_{ll}(a_{ll} - c_{ll} - b_{ll}) + b_{ll} + \delta(1-\theta)\pi c_{ll}\} \geq 0 \text{ and } e_{ll} \in [0,1]$$

$e_{ll,1} =$

$$
\begin{aligned}
&= 1 \text{ } \textit{if } 0 < a_{ll} - c_{ll} - b_{ll}, -(\frac{b_{ll} + \delta(1-\theta)\pi c_{ll}}{a_{ll} - c_{ll} - b_{ll}}) \leq 1 && \textit{(ll-1)} \\
&= 0 \text{ } \textit{if } 0 < a_{ll} - c_{ll} - b_{ll}, 1 < -(\frac{b_{ll} + \delta(1-\theta)\pi c_{ll}}{a_{ll} - c_{ll} - b_{ll}}) && \textit{(ll-2)} \\
&= 1 \text{ } \textit{if } 0 = a_{ll} - c_{ll} - b_{ll}, 0 \leq b_{ll} + \delta(1-\theta)\pi c_{ll} && \textit{(ll-3)} \\
&= 0 \text{ } \textit{if } 0 = a_{ll} - c_{ll} - b_{ll}, b_{ll} + \delta(1-\theta)\pi c_{ll} < 0 && \textit{(ll-4)} \\
&= \min\{\max\{0, -(\frac{b_{ll} + \delta(1-\theta)\pi c_{ll}}{a_{ll} - c_{ll} - b_{ll}})\}, 1\} \text{ } \textit{if } a_{ll} - c_{ll} - b_{ll} < 0. && \textit{(ll-5)}
\end{aligned}
$$

Consequently, we can derive the expected payoff values as follows;

$$
\begin{aligned}
V_{hh}((e_{hh,1}, 0), (C, T)) &= \frac{(1 - \delta(1-\pi)) r_{hh}(e_{hh,1}, e_{hh,1})}{(1-\delta)(1 - \delta + \delta\theta + \delta\pi - \delta\theta\pi)}, \\
V_{h}((e_{hh,1}, 0), (C, T)) &= \frac{\pi r_{hh}(e_{hh,1}, e_{hh,1})}{(1-\delta)(1 - \delta + \delta\theta + \delta\pi - \delta\theta\pi)},
\end{aligned}
$$

$$
\begin{aligned}
V_{ll}((0, e_{ll,1}), (T, C)) &= \frac{(1 - \pi\delta) r_{ll}(e_{ll,1}, e_{ll,1})}{(1-\delta)(1 - \delta\pi + \delta\theta\pi)}, \\
V_{l}((0, e_{ll,1}), (T, C)) &= \frac{(1 - \pi) r_{ll}(e_{ll,1}, e_{ll,1})}{(1-\delta)(1 - \delta\pi + \delta\theta\pi)}.
\end{aligned}
$$

When i-type players cooperates only with j-type players *where* $\{i, j\} = \{H, L\}$. *Let*

$$e_{hl}\{e_{hl}(a_{hl} - c_{hl} - b_{hl}) + b_{hl} + \delta(1-\theta)\pi c_{hl}\} \geq 0 \ \textit{and}\ e_{hl} \in [0, 1]$$

and

$$e_{lh}\{e_{lh}(a_{lh} - c_{lh} - b_{lh}) + b_{lh} + \delta(1-\theta)(1-\pi)c_{lh}\} \geq 0 \ \textit{and}\ e_{lh} \in [0, 1].$$

Now, we can figure out the solution $e_{hl,2} = e_{lh,2}$ *to the problem 2) of the Remark (6) as follows;* $e_{hl} =$

$$= \ 1 \ \textit{if}\ 0 < a_{hl} - c_{hl} - b_{hl}, -(\frac{b_{hl} + \delta(1-\theta)\pi c_{hl}}{a_{hl} - c_{hl} - b_{hl}}) \leq 1 \qquad \textit{(hl-1)}$$

$$= \ 0 \ \textit{if}\ 0 < a_{hl} - c_{hl} - b_{hl}, 1 < -(\frac{b_{hl} + \delta(1-\theta)\pi c_{hl}}{a_{hl} - c_{hl} - b_{hl}}) \qquad \textit{(hl-2)}$$

$$= \ 1 \ \textit{if}\ 0 = a_{hl} - c_{hl} - b_{hl}, 0 \leq b_{hl} + \delta(1-\theta)\pi c_{hl} \qquad \textit{(hl-3)}$$

$$= \ 0 \ \textit{if}\ 0 = a_{hl} - c_{hl} - b_{hl}, b_{hl} + \delta(1-\theta)\pi c_{hl} < 0 \qquad \textit{(hl-4)}$$

$$= \ \min\{\max\{0, -(\frac{b_{hl} + \delta(1-\theta)\pi c_{hl}}{a_{hl} - c_{hl} - b_{hl}})\}, 1\} \ \textit{if}\ a_{hl} - c_{hl} - b_{hl} < 0. \qquad \textit{(hl-5)}$$

$e_{lh} =$

$$
\begin{aligned}
&= 1 \ \textit{if}\ 0 < a_{lh} - c_{lh} - b_{lh}, -(\frac{b_{lh} + \delta(1-\theta)(1-\pi)c_{lh}}{a_{lh} - c_{lh} - b_{lh}}) \leq 1 && \textit{(lh-1)} \\
&= 0 \ \textit{if}\ 0 < a_{lh} - c_{lh} - b_{lh}, 1 < -(\frac{b_{lh} + \delta(1-\theta)(1-\pi)c_{lh}}{a_{lh} - c_{lh} - b_{lh}}) && \textit{(lh-2)} \\
&= 1 \ \textit{if}\ 0 = a_{lh} - c_{lh} - b_{lh}, 0 \leq b_{lh} + \delta(1-\theta)(1-\pi)c_{lh} && \textit{(lh-3)} \\
&= 0 \ \textit{if}\ 0 = a_{lh} - c_{lh} - b_{lh}, b_{lh} + \delta(1-\theta)(1-\pi)c_{lh} < 0 && \textit{(lh-4)} \\
&= \min\{\max\{0, -(\frac{b_{lh} + \delta(1-\theta)(1-\pi)c_{lh}}{a_{lh} - c_{lh} - b_{lh}})\}, 1\} \ \textit{if}\ a_{lh} - c_{lh} - b_{lh} < 0. && \textit{(lh-5)}
\end{aligned}
$$

$$e_{hl,2} = e_{lh,2} \equiv \min\{e_{hl}, e_{lh}\}.$$

And the expected payoff values are as follows;

$$
\begin{aligned}
V_{hl}((0, e_{hl,2}), (T, C)) &= \frac{(1-\pi\delta)r_{hl}(e_{hl,2}, e_{hl,2})}{(1-\delta)(1-\delta\pi+\delta\theta\pi)}, \\
V_{h}((0, e_{hl,2}), (T, C)) &= \frac{(1-\pi)r_{hl}(e_{hl,2}, e_{hl,2})}{(1-\delta)(1-\delta\pi+\delta\theta\pi)},
\end{aligned}
$$

$$
\begin{aligned}
V_{lh}((e_{lh,2}, 0), (C, T)) &= \frac{(1-\delta(1-\pi))r_{lh}(e_{lh,2}, e_{lh,2})}{(1-\delta)(1-\delta+\delta\theta+\delta\pi-\delta\theta\pi)}, \\
V_{l}((e_{lh,2}, 0), (C, T)) &= \frac{\pi r_{lh}(e_{lh,2}, e_{lh,2})}{(1-\delta)(1-\delta+\delta\theta+\delta\pi-\delta\theta\pi)}.
\end{aligned}
$$

Claim 56. *Suppose we found the solution* $(E^*, K^*) = ((e_{hh}, e_{hl}, e_{ll}), (k_{hh}, k_{hl}, k_{ll}))$ *to the optimization problem 1) or 2) in Remark (6) with* $V_{ij}^{d(1)}(E_i, K_i) = V_i(E_i, K_i)$ *where* $i, j \in \{h, l\}$. *Let* $e_{ij,t}$ *and* $e'_{ij,t}$ *denote an i-type deviator's action and a j-type punisher's action respectively at* t^{th} *stage of the total* T_{ij} *stages. If* $i = j \in \{h, l\}$ *or* $c_{hl} + b_{hl} - a_{hl} \geq 0$, *then there exists an action profile* $\{(e_{ij,t}, e'_{ji,t})\}_{t=1}^{T_{ii}}$ *for a punishment phase which can satisfy the incentive constraint and the participation constraint and is consistent with the assumption* $V_{ij}^{d(1)}(E_i, K_i) = V_i(E_i, K_i)$.

Proof. If we can find the action profile $\{(\alpha_{ij}^t, \beta_{ji}^t)\}_{t=1}^{T_{ij}} \in [0, 1]^{2T_{ij}}$ and punishment stage $T_{ij} \in \mathbb{N}$ such that

$$\begin{aligned}
1) V_i(E_i, K_i) &= r_{ij}(\alpha_{ij}^1, \beta_{ji}^1) + \delta(1-\theta) r_{ij}(\alpha_{ij}^2, \beta_{ji}^2) + \cdots \\
&\quad + \delta^{T_{ij}-1}(1-\theta)^{T_{ij}-1} r_{ij}(\alpha_{ij}^{T_{ij}}, \beta_{ji}^{T_{ij}}) \\
&\quad + \delta^{T_{ij}}(1-\theta)^{T_{ij}} V_{ij}(E_i, K_i) \\
&\quad + \frac{\delta\theta(1 - \delta^{T_{ij}}(1-\theta)^{T_{ij}})}{1 - \delta(1-\theta)} V_i(E_i, K_i),
\end{aligned}$$

$$
\begin{aligned}
2)V_j(E_j,K_j) \quad &\le \quad r_{ji}(\beta_{ji}^1,\alpha_{ij}^1)+\delta(1-\theta)r_{ji}(\beta_{ji}^2,\alpha_{ij}^2)+\cdots \\
&+\delta^{T_{ij}-1}(1-\theta)^{T_{ij}-1}r_{ji}(\beta_{ji}^{T_{ij}},\alpha_{ij}^{T_{ij}}) \\
&+\delta^{T_{ij}}(1-\theta)^{T_{ij}}V_{ji}(E_j,K_j) \\
&+\frac{\delta\theta(1-\delta^{T_{ij}}(1-\theta)^{T_{ij}})}{1-\delta(1-\theta)}V_j(E_j,K_j),
\end{aligned}
$$

$$
\begin{aligned}
3)r_{ij}(0,\beta_{ji})+\delta V_i(E_i,K_i) \quad &\le \quad r_{ij}(\alpha_{ij}^1,\beta_{ji}^1)+\delta(1-\theta)r_{ij}(\alpha_{ij}^2,\beta_{ji}^2)+\cdots \\
&+\delta^{T_{ij}-1}(1-\theta)^{T_{ij}-1}r_{ij}(\alpha_{ij}^{T_{ij}},\beta_{ji}^{T_{ij}}) \\
&+\delta^{T_{ij}}(1-\theta)^{T_{ij}}V_{ij}(E_i,K_i) \\
&+\frac{\delta\theta(1-\delta^{T_{ij}}(1-\theta)^{T_{ij}})}{1-\delta(1-\theta)}V_i(E_i,K_i),
\end{aligned}
$$

$$\cdots,$$

$$
\begin{aligned}
2+T_{ij})r_{ij}(0,\beta_{ji}^{T_{ij}})+\delta V_i(E_i,K_i) \quad &\le \quad r_{ij}(\alpha_{ij}^{T_{ij}},\beta_{ji}^{T_{ij}})+\delta(1-\theta)V_{ij}(E_i,K_i) \\
&+\delta\theta V_i(E_i,K_i),
\end{aligned}
$$

$$3+T_{ij})r_{ji}(0,\alpha_{ij})+\delta V_j(E_j,K_j) \leq (\beta^1_{ji},\alpha^1_{ij})+\delta(1-\theta)r_{ji}(\beta^2_{ji},\alpha^2_{ij})+\cdots$$
$$+\delta^{T_{ij}-1}(1-\theta)^{T_{ij}-1}r_{ji}(\beta^{T_{ij}}_{ji},\alpha^{T_{ij}}_{ij})$$
$$+\delta^{T_{ij}}(1-\theta)^{T_{ij}}V_{ji}(E_j,K_j)$$
$$+\frac{\delta\theta(1-\delta^{T_{ij}}(1-\theta)^{T_{ij}})}{1-\delta(1-\theta)}V_j(E_j,K_j)$$

$$\cdots,$$

$$2+2T_{ij})r_{ij}(0,\alpha^{T_{ij}}_{ij})+\delta V_i(E_i,K_i) \leq r_{ij}(\beta^{T_{ij}}_{ji},\alpha^{T_{ij}}_{ij})+\delta(1-\theta)V_{ij}(E_i,K_i)$$
$$+\delta\theta V_i(E_i,K_i),$$

, then we can choose $\{(\alpha^t_{ii},\beta^t_{ii})\}^{T_{ij}}_{t=1}$ as the actions of the punishment phase. Constraints 1)˜2) denote the participation constraint and constraints 3) ˜ $2+2T_{ii}$) denote the incentive constraint. Suppose $e_{ii}>0$ for $i=h$ or l. Then, it is true that $(k_{hh},k_{hl},k_{ll})=(C,T,C)$. So, we only need to figure out an action profile for the matches of the same type players. Since the same type of players are matched, constraint 1) means constraint 2) if $\alpha_{ii} \geq \beta_{ii}$. We can find out T_{ii} such that

$$\begin{aligned} V_i(E_i, K_i) &\leq r_{ii}(0,0) + \delta(1-\theta)r_{ii}(0,0) + \cdots \\ &+\delta^{T_{ii}-2}(1-\theta)^{T_{ii}-2}r_{ii}(0,0) + \delta^{T_{ii}-1}(1-\theta)^{T_{ii}-1}V_{ii}(E_i, \\ &K_i) + \frac{\delta\theta(1-\delta^{T_{ii}-1}(1-\theta)^{T_{ii}-1})}{1-\delta(1-\theta)}V_i(E_i, K_i) \\ \text{and } V_i(E_i, K_i) &\geq r_{ii}(0,0) + \delta(1-\theta)r_{ii}(0,0) + \cdots + \delta^{T_{ii}-1}(1-\theta)^{T_{ii}-1}r_{ii}(0,0) \\ &+\delta^{T_{ii}}(1-\theta)^{T_{ii}}V_{ii}(E_i, K_i) + \frac{\delta\theta(1-\delta^{T_{ii}}(1-\theta)^{T_{ii}})}{1-\delta(1-\theta)}V_i(E_i, K_i). \end{aligned}$$

Then, we can choose $(\alpha_{ii}, \alpha_{ii})$ such that

$$\begin{aligned} V_i(E_i, K_i) &= r_{ii}(0,0) + \delta(1-\theta)r_{ii}(0,0) + \cdots + \delta^{T_{ii}-1}(1-\theta)^{T_{ii}-1}r_{ii}(\alpha_{ii}, \alpha_{ii}) \\ &+\delta^{T_{ii}}(1-\theta)^{T_{ii}}V_{ii}(E_i, K_i) + \frac{\delta\theta(1-\delta^{T_{ii}}(1-\theta)^{T_{ii}})}{1-\delta(1-\theta)}V_i(E_i, K_i). \end{aligned}$$

Let $\{(e_{ii,t}, e'_{ii,t})\}_{t=1}^{T_{ii}}$ be $e_{ii,t} = e'_{ii,t} = 0$ if $t < T_{ii}$, $e_{ii,t} = e'_{ii,t} = \alpha_{ii}$ otherwise. Then, $\{(e_{ii,t}, e'_{ii,t})\}_{t=1}^{T_{ii}}$ satisfies constraints 1) ˜ $2 + 2T_{ii}$) above because $e_{ii,t} = e'_{ii,t} = 0$ if $t < T_{ii}$. Similarly, if $c_{hl} + b_{hl} - a_{hl} \geq 0$ then we can find action profile in the punishment phase when different type players cooperate. That is, we can find $(\alpha'_{ij}, \beta'_{ji}) \in [0,1]^2$ and $T_{ij} \in \mathbb{N}$ such that $\alpha'_{ij} > \beta'_{ji}$ and $\{(e_{ij,t}, e'_{ji,t})\}_{t=1}^{T'_{ij}}$ satisfies constraints 1) ˜ $2+2T$) above where $i \neq j$ and $e_{ij,t} = e'_{ji,t} = 0$ if $t < T'_{ij}$, $e_{ij,t} = \alpha'_{ij}$

and $e'_{ji,t} = \beta'_{ji}$ otherwise. If $c_{hl} + b_{hl} - a_{hl} \geq 0$, then

$$\begin{aligned} r_{ij}(0,0) + \delta V_i(E_i, K_i) &\leq r_{ij}(e_{ij}, 0) + \delta(1-\theta)V_{ij}(E_i, K_i) + \delta\theta V_i(E_i, K_i) \\ \text{and } r_{ji}(0, e_{ij}) + \delta V_j(E_j, K_j) &\leq r_{ji}(0, e_{ij}) + \delta(1-\theta)V_{ji}(E_j, K_j) + \delta\theta V_j(E_j, K_j) \end{aligned}$$

for $\{i, j\} = \{h, l\}$ where (E_i, K_i) is the solution to a problem in Remark (3) under the assumption $V_{ij}^{d(1)}(E_i, K_i) = V_i(E_i, K_i)$. So, A punisher does not need to hurt himself in order to punish the deviator. If we find out an action profile which satisfies constraints 1) and $2 + 2T$) such that $e_{ij,t} > e'_{ji,t}$ and $e'_{ji,t} = 0$ if $t < T_{ij}$, then it satisfies constraints 2) ˜ $2 + T$). It is because constraint 1) contains constraint 2) if $e_{ij,t} > e'_{ji,t}$ and it also means constraint 3)˜$2 + T$) if $e'_{ji,t} = 0$ for $t < T_{ij}$. We can always construct such profile which satisfies constraints 1) and 3)˜$2 + 2T$) such that $e_{ij,t} > e'_{ji,t}$ and $e'_{ji,t} = 0$ if $t < T_{ij}$. □

Claim 57. *Suppose $E^* = (e^*_{hh}, e^*_{hl}, e^*_{ll})$ satisfies condition 1) and 2) in Proposition (1) and $\{(e_{ij,t}, e'_{ji,t})\}_{t=1}^{T_{ij}}$ for $i, j \in \{h, l\}$ can be derived by the method in the proof of Claim (10) for a profile of actions in a punishment phase. Then, the following guides below can induce a social equilibrium; Let the punishment phase be a series of T_{ij} stages in which players in the pair play according to $\{(e_{ij,t}, e'_{ji,t})\}_{t=1}^{T_{ij}}$ for T_{ij} stages where $e_{ij,t}$ is the action for an i-type deviator and $e'_{ji,t}$ is the action of a j-type punisher at t^{th} stage of the total T_{ij} stages,* guide 1): play 'not Proposing a negotiation', 'Reject an offer', and 'Making an offer which contains $(e_{ij}, e_{ij}) = (1, 1)$ for the actions at the present match' in the negotiation process,

guide 2): play cooperation action according to E^* if players have not deviated from E^* in the match or if players have not deviated from E^* in the match after successfully finishing the punishment phase,

guide 3): play 'T' (=breaking up the relationship) if the cooperation level for the present match in E^* is *zero* and Play 'C' (=continuing the relationship) if the cooperation level for the present match in E^* is positive,

guide 4): play 'the actions in the punishment phase' if only i-type player deviated from "guide 2)" a stage ago in the match or if only i-type player deviated from $\{(e_{ij,t}, e'_{ji,t})\}_{t=1}^{T_{ij}}$ a stage ago during the punishment phase where $i, j \in \{h, l\}$,

guide 5): play 'according to the guide they were supposed to play a stage ago' if both players in the match deviated a stage ago.

Proof. These guides above can induce a social norm because it can be a complete contingent plan for each type of player and it is match-wise symmetric. It is a sequential equilibrium and satisfies negotiation-proofness by the definitions of $(E, K) = ((e^*_{hh}, e^*_{hl}, e^*_{ll}), (k^*_{hh}, k^*_{hl}, k^*_{ll}))$ and $\{(e_{ij,t}, e'_{ji,t})\}_{t=1}^{T_{ij}}$ for $i, j \in \{h, l\}$ and Proposition (1). $\square$

Proposition (2): If there exists a social equilibrium, then the profile of cooperation levels on the equilibrium path is unique.

Proof. Let $(E^1, K^1) = ((e^1_{hh}, e^1_{hl}, e^1_{ll}), (k^1_{hh}, k^1_{hl}, k^1_{ll}))$ and $(E^2, K^2) = ((e^2_{hh}, e^2_{hl}, e^2_{ll}), (k^2_{hh}, k^2_{hl}, k^2_{ll}))$ be profiles of actions on the equilibrium paths. By way of contradiction, we suppose $e^1_{hh} \neq e^2_{hh}$ or $e^1_{hl} \neq e^2_{hl}$. Let $(E^n_h, K^n_h) = ((e^n_{hh}, e^n_{hl}), (k^n_{hh}, k^n_{hl}))$ for $n \in \{1, 2\}$ be profiles of actions for H-payers on the equilibrium paths. Without loss of generality, we can assume $V_h(E^1_h, K^1_h) \geq V_h(E^2_h, K^2_h)$. If $V_h(E^2_h, K^2_h) > 0$, then according to Remark (6), the actions should be either $e^1_{hh}e^2_{hl} > 0$ or $e^1_{hl}e^2_{hh} > 0$ because players can cooperate only with one type of player and there are two kinds of matches for cooperative relationship. If $e^1_{hh}e^2_{hl} > 0$, then according to the incentive constraint and the condition 2) in Proposition (1),

$$\begin{aligned} r_{hl}(0, e^2_{hl}) + \delta V_h(E^2_h, K^2_h) &\leq \frac{r_{hl}(e^2_{hl}, e^2_{hl}) + \delta\theta V_h(E^2_h, K^2_h)}{1 - \delta(1-\theta)} \\ \text{and } r_{hh}(0, e) + \delta V_h(E^2_h, K^2_h) &> \frac{r_{hh}(e, e) + \delta\theta V_h(E^2_h, K^2_h)}{1 - \delta(1-\theta)} \end{aligned} \tag{1}$$

for all $e \in (0, 1]$. But, according to the incentive constraint of $(E^1_h, K^1_h) = ((e^1_{hh}, e^1_{hl}), (k^1_{hh}, k^1_{hl}))$,

$$r_{hh}(0, e^1_{hh}) + \delta V_h(E^1_h, K^1_h) \leq \frac{r_{hh}(e^1_{hh}, e^1_{hh}) + \delta\theta V_h(E^1_h, K^1_h)}{1 - \delta(1-\theta)}.$$

Since $V_h(E_h^1, K_h^1) \geq V_h(E_h^2, K_h^2)$, it is true that

$$(2) \qquad r_{hh}(0, e_{hh}^1) + \delta V_h(E_h^2, K_h^2) \leq \frac{r_{hh}(e_{hh}^1, e_{hh}^1) + \delta\theta V_h(E_h^2, K_h^2)}{1-\delta(1-\theta)}.$$

Since $e_{hh}^1 > 0$, the inequality (2) contradicts the inequality (1). Similarly we can find the contradiction in case of $e_{hl}^1 e_{hh}^2 > 0$. If $V_h(E_h^2, K_h^2) = 0$, then $e_{hh}^2 = e_{hl}^2 = 0$. According to the condition 1) and 2) in Proposition (1),

$$(3) \qquad r_{hl}(0, e) > \frac{r_{hl}(e, e)}{1-\delta(1-\theta)} \text{ and } r_{hh}(0, e) > \frac{r_{hh}(e, e)}{1-\delta(1-\theta)}$$

for all $e \in (0, 1]$. Since $\delta V > \frac{\delta\theta V}{1-\delta(1-\theta)}$ for all $V > 0$, we result in $e_{hh}^1 = e_{hl}^1 = 0$ from the inequalities (3) above. Therefore, we have shown that $e_{hh}^1 = e_{hh}^2$ and $e_{hl}^1 = e_{hl}^2$. With the same way, we can show that $e_{ll}^1 = e_{ll}^2$ and $e_{hl}^1 = e_{hl}^2$. Consequently, this result follows. □

Proposition (3): An Increase in $\delta(1-\theta)$ means non-decreases in cooperation levels on the equilibrium path in a social equilibrium. But, $\delta(1-\theta) \to 1$ does not mean $e_{ij} \to 1$ in a social equilibrium where $i, j \in \{h, l\}$.

Proof. According to the definitions of conditional equalities $ij-1$ ~ $ij-5$ in Remark (9) where $i, j \in \{h, l\}$, the first assertion follows. Let $c_{hh} = 11, a_{hh} = 10, b_{hh} = -1, c_{hl} = 2, a_{hl} = 1, b_{hl} = -1, c_{ll} = 1, a_{ll} = 0.3, b_{ll} = -0.2$, and $\pi = 2/3$.

When δ and θ are 0.9 and 0.01 respectively, $E^* = (e^*_{hh}, e^*_{hl}, e^*_{ll}) = (1, 0, 0.788)$ is derived in the Remark (6) with $V^{d(1)}_{ij}(E^*_i, K^*_i) = V_i(E^*_i, K^*_i)$ and satisfies two conditions in Proposition (1). Therefore, $E^* = (1, 0, 0.788)$ is the profile of cooperation levels on the equilibrium path in a social equilibrium. When δ and θ are between 0.9 and 1 and between 0 and 0.01 respectively, i.e. $\delta \in (0.9, 1)$ and $\theta \in (0, 0.01)$, we can find the profile of cooperation levels $E^* = (e^*_{hh}, e^*_{hl}, e^*_{ll})$ which is derived in the Remark (6) with $V^{d(1)}_{ij}(E^*_i, K^*_i) = V_i(E^*_i, K^*_i)$ and satisfies two conditions in Proposition (1). That is, there exists a social equilibrium if $\delta \in [0.9, 1)$ and $\theta \in (0, 0.01]$. When δ goes to 1 and θ goes to 0, the cooperation level for L-payers on the equilibrium path e^*_{ll} goes to $0.9\dot{3}$. This proves the second assertion. □

In this context, there are two different kinds of patience in the community. One is the discount factor of players δ and the other is the exogenous breaking possibility θ which determines a stability of the relationship. An increase in δ or a decrease in θ makes the long term cooperative relationship more valuable to players. Therefore, the players have more incentive to increase their cooperation levels. If the players increase their cooperation levels, then the payoffs of the deviators also increase. But, the players can not force the deviators to receive as worse severe punishment as the expected payoff in the pool of unmatched players, $V_i(E_i, K_i)$ for $i \in \{h, l\}$ because the deviators have the option to break up their relationship and can enter into the pool of unmatched players and there is no information flow across matches. As a result, players might not punish the deviators as severe as the deviators become worse off in the punishment phase than on the equilibrium path. In addition, the increase in δ or the decrease in θ also increases the expected payoff in the pool of unmatched players, $V_i(E_i, K_i)$. Consequently, the option to break up relationship blocks for the players to fully cooperate with their partner. We can easily see this by looking at the Incentive constraint;

$$r_{ij}(0, e_{ij}) + \delta V_i(E_i, K_i) \leq V_{ij}(E_i, K_i)$$

for $i, j \in \{h, l\}$. The players' cooperation levels are determined by the Incentive constraint. The increase in δ or the decrease in θ not only increases the expected payoff of the matched players on the right hand side on the inequality above, but also increases the expected payoff of the deviators in the punishment phase on the left hand side on the inequality. Therefore, the full cooperation level ,i.e. $e = 1$, may not be possible for some parameter values even though the players are patient enough and their relationship is sufficiently stable.

3.6. Feasibility of a steady state and Extension

3.6.1. Feasibility of a steady state

We have studied the state that π and θ are constant. Such a state is called a steady state. Here, we check the feasibility of the steady state.

Suppose that H-players cooperate only with H-players and L-players cooperate only with L-players, that is, the players cooperate only with the same type players. Let π'_h ,π'_l be the proportion of H-players and the proportion of L-players respectively who are cooperating with the same type player with respect to the total players. Let π^t_h ,π^t_l be the proportion of H-players and the proportion of L-players respectively in the pool of unmatched players at stage 't'. Then,

$\pi^t_h(1-\pi'_h-\pi'_l)$ denotes the proportion of unmatched H-players with respect to the total population at stage 't' and $(\pi^t_h-(\pi^t_h)^2)(1-\pi'_h-\pi'_l)+\theta\pi'_h$ denotes the proportion of unmatched H-players with respect to the total population at stage '$t+1$'.

$\pi^t_l(1-\pi'_h-\pi'_l)$ denotes the proportion of unmatched L-players with respect to the total population at stage 't' and $(\pi^t_l-(\pi^t_l)^2)(1-\pi'_h-\pi'_l)+\theta\pi^*_l$ denotes the proportion of unmatched L-players with respect to the total population at stage '$t+1$'.

If

$$(\pi_h^t)^2(1-\pi'_h-\pi'_l)=\theta\pi'_h \tag{1}$$

and

$$(\pi_l^t)^2(1-\pi'_h-\pi'_l)=\theta\pi'_l \tag{2}$$

, then they satisfy the sufficient condition for the steady state.

> **Claim:** For any given $\pi_h^t, \theta \in (0,1)$, there exist $\pi'_h, \pi'_l \in (0,1)$ subject to equations (1) and (2).

Proof. Fix π_h^t and θ. Suppose there exists π'_h and π'_l subject to (1) and (2). Then from (1) and (2),

$$\pi'_h=\pi'_l\frac{(\pi_h^t)^2}{(\pi_l^t)^2}. \tag{3}$$

From equations (1), (2) and (3), we can get

$$\begin{aligned}\pi'_h &= \frac{(\pi_h^t)^2}{\theta+(\pi_h^t)^2+(\pi_l^t)^2}\\ \pi'_l &= \frac{(\pi_l^t)^2}{\theta+(\pi_h^t)^2+(\pi_l^t)^2}.\end{aligned}$$

Obviously, $\pi'_h, \pi'_l \in (0,1)$. Therefore for any given $\pi^t_h, \theta \in (0,1)$, if we choose $\frac{(\pi^t_h)^2}{\theta+(\pi^t_h)^2+(\pi^t_l)^2}$ and $\frac{(\pi^t_l)^2}{\theta+(\pi^t_h)^2+(\pi^t_l)^2}$ as π'_h and π'_l respectively, then they satisfy equations (1) and (2) above. □

Similarly, we can check the feasibility of the steady state when H-players cooperate only with L-players and vice versa. Let π''_h ,π''_l be the proportion of H-players and the proportion of L-players respectively who are cooperating with a different type player with respect to the total players, then $\pi''_h = \pi''_l$. Therefore,

$\pi^t_h(1-2\pi''_h)$ denotes the proportion of unmatched H-players with respect to the total population at stage 't' and $(\pi^t_h - \pi^t_h\pi^t_l)(1-2\pi''_h) + \theta\pi''_h$ denotes the proportion of unmatched H-players with respect to the total population at stage '$t+1$'.

$\pi^t_l(1-2\pi''_h)$ denotes the proportion of unmatched L-players with respect to the total population at stage 't' and $(\pi^t_l - \pi^t_h\pi^t_l)(1-2\pi''_h) + \theta\pi''_h$ denotes the proportion of unmatched L-players with respect to the total population at stage '$t+1$'.

If

$$\pi^t_l\pi^t_h(1-2\pi''_h) = \theta\pi''_h = \theta\pi''_l \tag{4}$$

, then it satisfy the sufficient condition for the steady state.

Claim: For any given $\pi^t_h, \theta \in (0,1)$, there exist $\pi''_h, \pi''_l \in (0,1)$ subject to equation (4).

Proof. if we choose $\frac{\pi_l^t \pi_h^t}{\theta + 2\pi_l^t \pi_h^t} = \pi_h'' = \pi_l''$, then they satisfy equation (4) above. □

3.6.2. Extension

We have examined a complete information case. However, we can easily extend this model to an incomplete information case where players do not know their partner's type until they see their partner's action. In this incomplete information case, players' actions play a role of a signal with which players can figure out their partners' types. While in the traditional signaling games studied by Spence (1973), the signal itself reduces players' payoffs, in this model, however, the signal raises players' payoffs. So, players have an incentive to raise their signals. As a result, in H-H and L-L cooperation case, H-players have a bigger incentive to raise their signals because H-players' signals increase their payoffs more than L-players' signals do and those signals of H-players still satisfy incentive compatibility condition. In addition, we can enlarge our model from the two-type case to the several-type case. In those cases, we can find a variety of cooperation patterns. For example, players may cooperate only with a specific type of player or players cooperate with more than one type of players in equilibrium.

Finally, we have studied an extended version of the prisoner's dilemma game, because cooperation usually matters in the prisoner's dilemma game. But, we realize that the type-based payoff setting and the random-matching setting with the option of continuing the old relationship can be applied to a variety of games. For example, we can apply these settings to the battle of sexes game. We expect that it can provide another explanation of male-female matching.

3.7. Conclusion

We have studied the possibility of a long term cooperative relationship in an extended version of the prisoner's dilemma game with a continuum of players where a stage game is infinitely repeated between every two players who are randomly matched at each stage and they have the option to continue their relationship. The distinguishing characteristic of the model is an asymmetric player assumption. Players are characterized by their cooperation ability which has an effect on their partner's payoff as well as their own payoff. That is, players' payoff in the pair depends on the players' cooperation abilities in the pair as well as their cooperation actions. The difference between cooperation abilities among players allows players to distinguish their partner's type. Because of a low expected payoff through cooperation or a high expected payoff through deviation, the players may not cooperate with a specific type of player. Therefore, the match with the other type of player could be endowed with the scarcity value. This scarcity value makes the retaliation of breaking up the relationship fatal, where cooperation in the relationship is possible. The players can utilize the effective retaliation of terminating relationship to sustain the cooperation with the type of player.

We have seen that a social equilibrium must be unique if it exists. This is because of the negotiation-proofness. The negotiation-proofness requires a unique possible cooperation pattern and the best cooperation levels in the model according to which players cooperate on the equilibrium path.

Finally, we have examined whether or not a kind of Folk theorem can be applied to this model and found that it can not. Because a deviator has the option to break up his relationship and join in a new match with another partner and because a player's past action does not affect his future expected payoff, the players may not punish the deviator as severely as the deviator becomes worse off in the punishment phase than on the equilibrium path. This limitation of the punishment may block the players in fully cooperating with their partner.

References for the first chapter

(1) Aumann, Robert J. and Hart, Sergiu (2003): "Long Cheap Talk," *Econometrica*, 71, 1619–1660.

(2) Austen-Smith, David (1990): "Information Transmission in Debate," *American Journal of Political Science*, 34, 124–152.

(3) Baliga, Sandeep and Morris, Stephen (2002): "Co-ordination, Spillovers, and Cheap Talk," *Journal of Economic Theory*, 105, 450–468.

(4) Baliga, Sandeep and Sjöström, Tomas (2004): "Arms Races and Negotiations," *Review of Economic Studies*, 71, 351–369.

(5) Banerjee, Abhijit V. (1992): "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797–817.

(6) Banks, Jeffrey S. and Sobel, Joel (1987): "Equilibrium Selection in Signaling Games," *Econometrica*, 55, 647–661.

(7) Baron, David P. (2005): "Competing for the Public through the News Media," *Journal of Economics and Management Strategy*, 14, 339–376.

(8) Baron, David P. (2006): "Persistent Media Bias," *Journal of Public Economics*, 90, 1–36.

(9) Battaglini, Marco (2002): "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70, 1379–1401.

(10) Bernhardt, Dan; Krasa, Stefan and Polborn, Mattias (2006): "Political Polarization and the Electoral Effects of Media Bias," CESifo Working Paper Series No. 1798.

(11) Besley, Timothy and Prat, Andrea (2004): "Handcuffs for the Grabbing Hand?–Media Capture and Government Accountability," CEPR Discussion Paper No. 3132.

(12) Bhattacharya, Sudipto (1979): "Imperfect Information, Dividend Policy, and 'The Bird in the Hand' Fallacy," *Bell Journal of Economics*, 10, 259–270.

(13) Cho, In-Koo and Kreps, David M. (1987): "Signaling Games and Stable Equilibria," Q*uarterly Journal of Economics*, 102, 179–222.

(14) Crawford, Vincent P. and Sobel, Joel (1982): "Strategic Information Transmission," *Econometrica*, 50, 1431–1451.

(15) DeFleur, Melvin L. and Ball-Rokeach, Sandra (1989): "Theories of Mass Communication," Chapter 6–7, Longman Press, Fifth Edition.

(16) Dyck, Alexander and Zingales, Luigi (2003): "The Media and Asset Prices," mimeo.

(17) Farrell, Joseph (1993): "Meaning and Credibility in Cheap-talk Games," *Games and Economic Behavior*, 5, 514–531.

(18) Farrell, Joseph and Gibbons, Robert (1989a): "Cheap Talk with Two Audiences," *American Economic Review*, 79, 1214–1223.

(19) Farrell, Joseph and Gibbons, Robert (1989b): "Cheap Talk Can Matter in Bargaining," *Journal of Economic Theory*, 48, 221–237.

(20) Farrell, Joseph and Rabin, Matthew (1996): "Cheap Talk," *Journal of Economic Perspectives*, 10, 103–118.

(21) Fudenberg, Drew and Tirole, Jean (1991): "Perfect Bayesian Equilibrium and Sequential Equilibrium," *Journal of Economic Theory*, 53, 236–260.

(22) Gamson, William A.; Croteau, David; Hoynes, William and Sasson, Theodore (1992): "Media Images and the Social Construction of Reality," *Annual Review of Sociology*, 18, 373–393.

(23) Gentzkow, Matthew and Shapiro, Jesse M. (2006): "Media Bias and Reputation," *Journal of Political Economy*, 114, 280–316.

(24) Heyward, Andrew (2004): "CBS Statement on Bush Memos," CBSNews.com, September 20.

(25) Kartik, Navin; Ottaviani, Marco and Squintani, Francesco (2006): "Credulity, Lies, and Costly Talk," *Journal of Economic Theory*, Forthcoming.

(26) Kirkpatrick, David D. and Fabrikant, Geraldine (2003): "Advertisers and Wall Street Welcome Moves to Change," New York Times, June 6.

(27) Krishna, Vijay and Morgan, John (2001): "A Model of Expertise," Q*uarterly Journal of Economics*, 116, 747–775.

(28) Manelli, Alejandro M. (1997): "The Never-a-Weak-Best-Response Test in Infinite Signaling Games," *Journal of Economic Theory*, 74, 152–173.

(29) Milgrom, Paul R. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 380–391.

(30) Milgrom, Paul R. and Roberts, John (1986): "Price and Advertising Signals of Product Quality," *Journal of Political Economy*, 94, 796–821.

(31) Morris, Stephen and Shin, Hyun Song (2002): "Social Value of Public Information," *American Economic Review*, 92, 1521–1534.

(32) Mullainathan, Sendhil and Shleifer, Andrei (2005): "The Market for News," *American Economic Review*, 95, 1031–1053.

(33) Rather, Dan (2004): "Dan Rather Statement on Memos," CBSNews.com, September 20.

(34) Scharfstein, David S. and Stein, Jeremy C. (1990): "Herd Behavior and Investment," *American Economic Review*, 80, 465–479.

(35) Schelling, Thomas C. (1960): "The Strategy of Conflict," Chapter 9, Harvard University Press.

(36) Seidmann, Daniel J. and Winter, Eyal (1997): "Strategic Information Transmission with Verifiable Messages," *Econometrica*, 65, 163–169.

(37) Shirer, William L. (1990): "The Rise and Fall of the Third Reich," Chapter 16, MJF Books Press.

(38) Sobel, Joel (1985): "A Theory of Credibility," *Review of Economic Studies*, 52, 557–573.

(39) Spence, Michael (1973): "Job Market Signaling," Q*uarterly Journal of Economics*, 87, 355–379.

(40) Stein, Jeremy C. (1989): "Cheap Talk and the Fed: A Theory of Imprecise Policy Announcements," *American Economic Review*, 79, 32–42.

(41) Strömberg, David (2004): "Mass Media Competition, Political Competition, and Public Policy," *Review of Economic Studies*, 71, 265–284.

(42) Van Damme, Eric (1989): "Stable Equilibria and Forward Induction," *Journal of Economic Theory*, 48, 476–496.

References for the second chapter

(1) J. Hirshleifer, Anarchy and its breakdown, J. Polit. Economy 103, 1995, 26-52.

(2) J. Hirshleifer, The paradox of power, Econ. Politics 3, 1991, 177-200.

(3) J. Jordan, Pillage and Property, J. Econ. Theory, 2005, forthcoming.

(4) K. Konrad and S. Skaperdas, Extortion, Economica 65, 1998, 461-77.

(5) W. Lucas,Von Neumann–Morgenstern stable sets, in: R. Aumann, S. Hart (Eds.), Handbook of Game Theory, Elsevier, Amsterdam, 1992, 543-90.

(6) A. Muthoo, A model of the origins of basic property rights, Games Econ. Behav. 3, 1991, 177-200

(7) M. Piccione and A. Rubinstein, Equilibrium in the jungle, 2006, mimeo.

(8) A. Roth, Subsolutions and the supercore of cooperative games, Math. Operations Res. 1, 1976, 43-49.

(9) S. Skaperdas, Cooperation, conflict and power in the absence of property rights, Amer. Econ. Rev. 82, 1992, 720-739.

References for the third chapter

(1) Abreu, D (1988), "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383-396.

(2) Datta, Saikat (1996), "Building Trust," (Mimeo).

(3) Ellison, G (1994), "Cooperation in the prisoner's dilemma with anonymous random matching," *Review of Economic Studies*, 61, 567-588.

(4) Fudenberg, D and Maskin, E (1986), " The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54, 533-556.

(5) Fudenberg, D and Tirole J (1991), "Game Theory," MIT Press, Cambridge.

(6) Ghosh, P and Ray, D (1996), "Cooperation in Community Interaction without Information flows," *Review of Economic Studies*, 63, 491-519.

(7) Kandori, M (1992), "Social norms and community enforcement," *Review of Economic Studies*, 59, 63-80.

(8) Mattuck, A (1999), "Introduction to Analysis," Prentice-Hall, Inc.

(9) Spence, M. (1973), "Job Market Signaling,? Quarterly Journal of Economics, 87, 355-379.

(10) Watson, Joel (1999), "Starting small and Commitment," (Mimeo).

Vita of Michael Jung

Name: Han Joon (Michael) Jung

CITIZENSHIP: Republic of Korea, F-1 Visa

EDUCATION: Ph.D., Economics, Pennsylvania State University, expected May 2007

B.A., Economics and History, Sung-Kyun-Gwan University, Korea, 2001

Ph.D. THESIS: "Essays in Game Theory and Political Economy"

Thesis Advisor: Professor James Jordan

FIELDS: Primary: Microeconomic Theory, Game Theory, Political Economy

Secondary: Industrial Organization

TEACHING EXPERIENCE: Teaching Assistant: Introductory Microeconomics (5 semesters), Introductory Macroeconomics (2 semester), Game Theory (2 semesters), Monetary Theory (1 semester)

RESEARCH EXPERIENCE: Research Assistant for James Jordan (Spring-Summer 2005, Summer 2006)

PRESENTATION: "Strategic Information Transmission through the Media" Mid-west Economic Theory Meeting, October 2006.