This is a repository copy of *Reject and Renegotiate : the Shapley value in multilateral bargaining*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/75036/

Version: Submitted Version

Article:
Ju, Yuan orcid.org/0000-0002-7541-9856 (2012) Reject and Renegotiate : the Shapley value in multilateral bargaining. Journal of Mathematical Economics. pp. 431-436. ISSN 0304-4068

https://doi.org/10.1016/j.jmateco.2012.08.009

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Accepted Manuscript

Reject and renegotiate: The Shapley value in multilateral bargaining

Yuan Ju

PII: S0304-4068(12)00064-X
DOI: 10.1016/j.jmateco.2012.08.009
Reference: MATECO 1745

To appear in: *Journal of Mathematical Economics*

Received date: 16 May 2012
Revised date: 10 August 2012
Accepted date: 28 August 2012

Please cite this article as: Ju, Y., Reject and renegotiate: The Shapley value in multilateral bargaining. *Journal of Mathematical Economics* (2012), doi:10.1016/j.jmateco.2012.08.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Reject and renegotiate: the Shapley value in multilateral bargaining*

Yuan Ju†

September 6, 2012

*I am grateful to the associate editor and an anonymous reviewer for the valuable suggestions as well as the detailed and constructive comments. Helpful discussions with the participants at the 2012 RCGEB Summer Workshop in Qingdao, China, and the 4th World Congress of the Game Theory Society in Istanbul, Turkey are appreciated. Of course, any remaining error is mine. The financial support of the Super Pump Priming Fund from the University of York is gratefully acknowledged.

†Department of Economics and Related Studies, University of York, Heslington, York, YO10 5DD, UK. Tel: +44-1904-324678, Fax: +44-1904-323759, E-mail: yuan.ju@york.ac.uk.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Abstract

This paper investigates three distinctive and intuitive renegotiation bargaining protocols that all yield the Shapley value as the unique subgame perfect equilibrium outcome. These protocols, built on the multi-bidding procedure of Pérez-Castrillo and Wettstein (2001), allow more freedom in multilateral bargaining where rejected players can further negotiate and form coalitions. The self-duality of the Shapley value plays a key role in the second and third results. Moreover, these renegotiation protocols allow an actual play along the equilibrium path to restore the Shapley value in case of a 'mistake' made before.

Keywords: bargaining; subgame perfect equilibrium; Shapley value; renegotiation.

JEL code: C71; C72; D62

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1 Introduction

Almost at the same time when the two fundamental solution concepts for game theory, Nash equilibrium (Nash (1950a)) and Nash bargaining solution (Nash (1950b)), were established, John Nash (1953) pointed out the importance and necessity of using both axiomatic and strategic approaches to bridge the gap between its non-cooperative side and cooperative side. A comprehensive survey on this research agenda, called Nash program, is provided by Serrano (2005), and broadly for implementation theory, we refer to Jackson (2001). Discussions on this program in relation to mechanism design theory can be found in Trockel (2002).[1] In his presidential address to the Econometric Society, Maskin (2003) outlined coalitions and externalities as core topics for game theory in the coming years and highlighted adopting both axiomatic and strategic approaches, which was reiterated in his talk at the Nobel Panel meeting of the 2008 Games Congress in Chicago.

Although the Shapley value, a major solution concept for cooperative games, was invented in the same period (Shapley (1953)) along the work of Nash, it was not until over thirty years later that the first non-cooperative study on this solution concept was proposed: a bilateral bargaining procedure by Gul (1989, 1999). Hart and Mas-Collel (1996) constructed a multilateral bargaining procedure to obtain the Shapley value. This procedure is further studied in Krishna and Serrano (1995). By endogenizing the selection of a proposer, Pérez-Castrillo and Wettstein (2001) developed the multi-bidding mechanism (Pérez-Castrillo and Wettstein (2002)) that yields the Shapley value in actual terms. Pérez-Castrillo and Wettstein (2001) also offered an extensive discussion of the non-cooperative study on the Shapley value.

The current research contributes to the non-cooperative study of the multilateral coalitional bargaining in two main aspects.

Firstly, the option for rejected players to have chance to form a coalition again and then renegotiate with existing players is modeled and analyzed. So far in the literature most bargaining protocols for the Shapley value require a rejected proposer to simply leave the game and stand alone, i.e., losing all the possibilities of forming coalitions with anyone else. This contradicts our real life observation that those who are excluded from a party (or a certain organization) may well organize into a new party, and even further renegotiate with the old party for potential greater benefits. Such a possibility seems necessary to explore. Based on the original multi-bidding mechanism proposed by Pérez-Castrillo and Wettstein (2001), we construct three non-cooperative bargaining protocols to study the option of renegotiation.

[1] It is worthwhile to note an interesting research by Sun, Trockel and Yang (2008) that adds coalition formation issue into the broad Nash program, where a cooperative solution concept is supported by competitive outcomes of a decentralized production economy.

ACCEPTED MANUSCRIPT

These renegotiation bargaining protocols not only seem realistic, but also lead to actual plays on equilibrium paths. Unlike most bargaining procedures in the literature that only obtain the Shapley value at the first stage in subgame perfect equilibrium (SPE), the renegotiation protocols introduced in the paper allow players to actually go through the following stages and still realize the Shapley value.

Secondly, the paper offers a robustness study of the Shapley value in non-cooperative bargaining with different renegotiation protocols. Ju and Wettstein (2009) proposes a generalized bidding mechanism and shows that by introducing a renegotiation stage[2] a bargaining game may generate completely different value allocations in equilibrium. The three renegotiation protocols presented in this paper restore the Shapley value in SPE, despite the expanded options in bargaining, which provide additional support to this major cooperative solution concept.

In addition, it is worthwhile to note that the self-duality of the Shapley value, a desired axiomatic property but under-explored in the literature of Nash program, plays an important role in the proofs of this study. Hence, the non-cooperative approach further enhances our understanding of the cooperative side of the Shapley value.

The next section provides the preliminaries. In Section 3 we construct the three distinctive non-cooperative bargaining protocols with renegotiation and show that they all yield the Shapley value in SPE. The final section offers some concluding remarks.

2 The cooperative model and the Shapley value

We denote by $N = \{1, ..., n\}$ the set of players, and by $S \subseteq N$ a coalition of players. A transferable utility (TU) game in characteristic form is denoted by (N, v) where $v : 2^N \to \mathbb{R}$ is a characteristic function satisfying $v(\emptyset) = 0$. The class of all TU games with player set N is denoted by TU^N. Throughout the paper, $|S|$ denotes the cardinality of S, and for simplicity we use $|S| = s$. When no confusion arises, let $|N| = n$. For a coalition S, $v(S)$ is the total payoff that the members in S can obtain if S forms. For notational simplicity, given $i \in N$, we use $v(i)$ instead of $v(\{i\})$ to denote the stand-alone payoff of player i. A *value* is a mapping f which associates with every game (N, v) a vector in $\mathbb{R}^N$. A value determines the payoffs for every player in the game.

Given a cooperative game (N, v) and a subset $S \subseteq N$, we define the subgame $(S, v|_S)$ by assigning the value $v|_S(T) \equiv v(T)$ for any $T \subseteq S$.

[2]That renegotiation is restricted for each rejected player and the existing agreed coalition only, but does not allow rejected players to form coalition by themselves, whereas it is permissible in the current paper.

ACCEPTED MANUSCRIPT

We denote by ϕ the Shapley value for game (N, v) which is defined by

$$\phi_i(N, v) = \sum_{S \subseteq N \setminus \{i\}} \frac{s!(n-s-1)!}{n!} [v(S \cup \{i\}) - v(S)]$$

for all $i \in N$. It is the unique value that satisfies efficiency, additivity, symmetry and the null player property.

Moreover, it is well-known that the Shapley value is self-dual (cf. Kalai and Samet (1987)). Given a game $v \in TU^N$, its dual game (N, v^d) is defined by $v^d(S) = v(N) - v(N \setminus S)$ for all $S \subseteq N$. A solution concept f satisfies self-duality if $f(v) = f(v^d)$ for every $v \in TU^N$.

3 The renegotiation mechanisms

In this section, we construct three different bargaining protocols with the option of renegotiation and show that all of them yield the Shapley value as the unique SPE outcome. Although these three mechanisms are different in the specific renegotiation design, they share the same basic structure, which can be described briefly and informally as follows.

Players firstly participate in a bidding stage *a la* Pérez-Castrillo and Wettstein (2001) to choose a proposer.[3] At the next stage the proposer offers a vector of payments to all other players in exchange for joining her to form the grand coalition. This is equivalent to saying that the proposer makes a scheme showing how to split $v(N)$ among all the players. At stage 3, all the other players sequentially decide to agree or reject the offer. The offer is accepted if all players accept it, and is rejected if at least one player rejects. In case of acceptance the grand coalition forms and the proposer receives $v(N)$ out of which she pays out the offers made. In case of rejection this (rejected) proposer 'waits' (instead of getting her stand-alone payoff) while all the other players go again through the same game.

We now investigate the possibility that the rejected proposers can further bargain among themselves to form a coalition, following which they may be able to *renegotiate* with the 'incumbent' players, to obtain what we term the *renegotiation mechanisms*.

Inspired by the real world cases such as a rejected party leader can form a new party with his followers, or a departed founding member of a company may establish a new firm by hiring some staff from the former company, we construct the first renegotiation mechanism. It specifies that the first rejected proposer has power to unite all rejected proposers and then renegotiate with the incumbent players. The second renegotiation mechanism is a natural variant, which has the incumbent player make a renegotiation offer to the first rejected proposer. The third mechanism is more flexible. Rather than

[3] As can be seen from the following analysis, the reason we build the renegotiation bargaining protocols on this bidding stage is that it helps to endogenously select representatives of incumbent players, which facilitates the renegotiation in a natural way.

ACCEPTED MANUSCRIPT

having the first rejected proposer have power to unite all rejected proposers, it allows every proposer, when rejected, to be able to bargain and form a coalition with the immediately preceding rejected proposer.

As for how these three mechanisms precisely proceed and differ in the respective renegotiation stages, we refer to the following formal descriptions.

Mechanism A. If there is only one player $\{i\}$, she simply receives $v(i)$. When the player set $N = \{1, \ldots, n\}$ consists of two or more players, the mechanism is defined recursively.

Stages 1 to 3 provide for any set of (active) players S a proposer in S, chosen via a bidding procedure (stage 1), an offer made by the proposer to the rest of the players in S (stage 2) who then accept or reject (stage 3). If stage 3 ends with a rejection, all players in S other than the rejected proposer proceed again through stages 1 to 3 where the set of active players is reduced by excluding the rejected proposer. If stage 3 ends with acceptance, for $S = N$ the game ends; but for a coalition S smaller than N, the game moves to stage 4. At stage 4, all previously rejected proposers, i.e., the inactive players $N \backslash S$, have one more chance to organize themselves. The earliest rejected proposer makes a take-it-or-leave-it offer to every other player in $N \backslash S$. If this offer is rejected, then the game ends with players in $N \backslash S$ receiving their stand-alone payoffs. If the offer is accepted, then the game proceeds to stage 5, where the earliest rejected proposer, as now the representative of $N \backslash S$, makes a 'renegotiation' offer to the proposer of S. If this renegotiation offer is accepted, then S and $N \backslash S$ merge into the grand coalition and the game ends; otherwise, the game stops such that the two sides stay split.

The mechanism starts with $S = N$.

Stage 1: Each player $i \in S$ makes $s-1$ bids $b^i_j \in \mathbb{R}$ with $j \in S \backslash \{i\}$. For each $i \in S$, define the *net bid* of player i by $B^i = \sum_{j \in S \backslash \{i\}} b^i_j - \sum_{j \in S \backslash \{i\}} b^j_i$. The net bids are used to measure players' willingness to become the proposer. Let $i_s = \text{argmax}_{i \in S}(B^i)$ where in case of a non-unique maximizer we choose any of these maximal bidders to be the 'winner' with equal probability. Once the winner i_s has been chosen, player i_s pays every player $j \in S \backslash \{i_s\}$ her bid $b^{i_s}_j$.

Stage 2: Player i_s makes a vector of offers $x^{i_s}_j \in \mathbb{R}$ to every player $j \in S \backslash \{i_s\}$. (This offer is additional to the bids paid at stage 1.)

Stage 3: The players in S other than i_s, sequentially, either accept or reject the offer. If at least one player rejects it, then the offer is rejected. Otherwise, the offer is accepted.

ACCEPTED MANUSCRIPT

If the offer made at stage 2 is rejected, all players in S other than i_s proceed again through the mechanism from stage 1 where the set of active players is $S\backslash\{i_s\}$.[4] Meanwhile, player i_s, together with all previously rejected proposers $N\backslash S$, waits for the negotiation outcome of $S\backslash\{i_s\}$.

If the offer of i_s is accepted, we have to distinguish between two cases where $S = N$ and $S \neq N$. In the case where $S = N$, which means that all players agree with the proposer on the scheme of sharing $v(N)$, the game ENDS. Then, the final payoff to player $j \neq i_n$ is $x_j^{i_n} + b_j^{i_n}$ while player i_n receives $v(N) - \sum_{j\in N\backslash\{i_n\}} x_j^{i_n} - \sum_{j\in N\backslash\{i_n\}} b_j^{i_n}$. In the case where $S \neq N$, the game moves to stage 4.

Stage 4: The first rejected proposer i_n proposes a vector of take-it-or-leave-it payments $y_{i_k}^{i_n} \in \mathbb{R}$ to every rejected proposer i_k where $k = s+1, ..., n-1$. These players, sequentially, either accept or reject the proposal. Again, acceptance requires unanimity.

If the proposal is rejected, then the game ends with all rejected proposers in $N\backslash S$ standing alone. Thus, i_s finally receives $v(S) - \sum_{j\in S\backslash\{i_s\}} b_j^{i_s} - \sum_{j\in S\backslash\{i_s\}} x_j^{i_s} + \sum_{k=s+1}^{n} b_{i_s}^{i_k}$, each player $j \in S\backslash\{i_s\}$ finally receives $\sum_{k=s}^{n} b_j^{i_k} + x_j^{i_s}$, and each player i_k for $k = s+1, ..., n$ finally receives $v(i_k) - \sum_{j\in N\backslash\{i_k,i_{k+1},...,i_n\}} b_j^{i_k} + \sum_{l\in\{i_{k+1},...,i_n\}} b_{i_k}^{l}$.

If the proposal is accepted, then i_n pays the offer $y_{i_k}^{i_n} \in \mathbb{R}$ to every player i_k where $k = s+1, ..., n-1$, and the coalition $N\backslash S$ is formed, of which i_n becomes the representative. The game proceeds to stage 5.

Stage 5: Player i_n makes a renegotiation offer $z_{i_s}^{i_n} \in \mathbb{R}$ to i_s who is the representative of S.

If the renegotiation offer is rejected by i_s, then the game ends with both sides, S and $N\backslash S$, staying apart. Thus, i_s finally receives $v(S) - \sum_{j\in S\backslash\{i_s\}} b_j^{i_s} - \sum_{j\in S\backslash\{i_s\}} x_j^{i_s} + \sum_{k=s+1}^{n} b_{i_s}^{i_k}$, each player $j \in S\backslash\{i_s\}$ finally receives $\sum_{k=s}^{n} b_j^{i_k} + x_j^{i_s}$, while i_n finally receives $v(N\backslash S) - \sum_{j\in N\backslash\{i_n\}} b_j^{i_n} - \sum_{k=s+1}^{n-1} y_{i_k}^{i_n}$ and each player i_k for $k = s+1, ..., n-1$ finally receives $y_{i_k}^{i_n} - \sum_{j\in N\backslash\{i_k,i_{k+1},...,i_n\}} b_j^{i_k} + \sum_{l\in\{i_{k+1},...,i_n\}} b_{i_k}^{l}$.

If i_s accepts the renegotiation offer, then i_n pays $z_{i_s}^{i_n}$ to i_s so that the grand coalition N is formed. Hence, the game ends and i_s finally receives $z_{i_s}^{i_n} - \sum_{j\in S\backslash\{i_s\}} b_j^{i_s} - \sum_{j\in S\backslash\{i_s\}} x_j^{i_s} + \sum_{k=s+1}^{n} b_{i_s}^{i_k}$, each player $j \in S\backslash\{i_s\}$ finally receives $\sum_{k=s}^{n} b_j^{i_k} + x_j^{i_s}$, while i_n finally receives $v(N) - z_{i_s}^{i_n} - \sum_{j\in N\backslash\{i_n\}} b_j^{i_n} - \sum_{k=s+1}^{n-1} y_{i_k}^{i_n}$ and each player i_k for $k = s+1, ..., n-1$ finally receives $y_{i_k}^{i_n} - \sum_{j\in N\backslash\{i_k,i_{k+1},...,i_n\}} b_j^{i_k} + \sum_{l\in\{i_{k+1},...,i_n\}} b_{i_k}^{l}$.

[4] To make it clearer, here we explicitly explain the amount of payoff that players in $S\backslash\{i_s\}$ will bargain for, although it is incorporated in the description of the following stages provided below. Because of the chance of renegotiation at stages 4 and 5, players in $S\backslash\{i_s\}$ bid for becoming the proposer i_{s-1} so as to win the offer $z_{i_{s-1}}^{i_n}$ made by i_n at stage 5. In equilibrium, it equals $v(S\backslash\{i_s\})$.

ACCEPTED MANUSCRIPT

We note that in the case the mechanism reaches the situation where the set of active players consists of one player only, i.e., $|S| = 1$, the corresponding stages 1 to 3 are redundant and this single player is considered as the proposer for herself whose offer is accepted immediately and the game continues to stage 4.

The idea of having the first rejected proposer making a take-it-or-leave-it offer to all other rejected players is consistent with many real world observations. After a CEO quits a company, his followers may go with him, and then he may set up a competing firm with these colleagues. And it is not rare that these two firms could well merge some time later. Similar stories apply to political parties and their leaders.

We will show that for any (strictly) superadditive[5] game (N, v) (i.e., $v(S \cup T) > v(S) + v(T)$ for all $S, T \subseteq N$ with $S \cap T = \emptyset$), all SPE of Mechanism A yield the same outcome, which coincides with the payoff vector $\phi(N, v)$ as prescribed by the Shapley value.

Theorem 3.1 *For any superadditive game (N, v), the subgame perfect equilibrium outcomes of Mechanism A coincide with the payoff vector $\phi(N, v)$ as prescribed by the Shapley value.*

To facilitate the understanding of the result, we sketch a key reasoning that underlies the feature of the mechanism. One readily sees that, among rejected proposers, only i_n can potentially get extra payoff beyond her stand-alone payoff, due to the right to propose to all the rejected players at stage 4. In SPE, she will offer the stand-alone payoffs to these players. Anticipating this, any subsequent proposer would not make an offer that will be rejected: by superadditivity, we know that making an offer that will be accepted results in a higher payoff to the proposer than his stand-alone payoff, the SPE outcome if being rejected. Hence, reasoning further backward, the first proposer i_n, if her offer is rejected, will actually end up with her stand-alone payoff. Consequently, despite a potential advantage of gaining a big surplus by making a take-it-or-leave-it offer to those who got rejected after her, it is impossible (because everyone else is smart) for the first rejected proposer to substantiate it. With this in mind and due to superadditivity, in equilibrium, i_n will also make an offer that is accepted by all players in $N \backslash \{i_n\}$.

Proof (of Theorem 3.1)
Let (N, v) be a superadditive game. Before showing that every SPE yields the payoff vector $\phi(N, v)$, we first show that the Shapley value is an SPE outcome by explicitly constructing

[5] A weaker condition called strict zero-monotonicity (i.e., $v(S) > v(S \backslash \{i\}) + v(\{i\})$ for all $S \subseteq N$ and $i \in S$) is sufficient for Theorem 3.1. We use superadditivity here to be consistent in presenting all the results as it is required for Theorem 3.2 and Theorem 3.4.

ACCEPTED MANUSCRIPT

such an SPE. Consider the following strategies (we describe it for the active set of players, S):

At stage 1, each player $i \in S$ announces $b^i_j = \phi_j(S, v|_S) - \phi_j(S\backslash\{i\}, v|_{S\backslash\{i\}})$ for every $j \in S\backslash\{i\}$.

At stage 2, a proposer, player i_s, offers $x^{i_s}_j = \phi_j(S\backslash\{i_s\}, v|_{S\backslash\{i_s\}})$ to every $j \in S\backslash\{i_s\}$.

At stage 3, any player $j \in S\backslash\{i_s\}$ accepts any offer which is greater than or equal to $\phi_j(S\backslash\{i_s\}, v|_{S\backslash\{i_s\}})$ and rejects any offer strictly less than $\phi_j(S\backslash\{i_s\}, v|_{S\backslash\{i_s\}})$.

At stage 4, the first rejected proposer i_n makes an offer $y^{i_n}_{i_k} = v(\{i_k\})$ to any subsequently rejected proposer i_k where $k = s+1, ..., n-1$. And any i_k, where $k = s+1, ..., n-1$, accepts an offer $y^{i_n}_{i_k} \geq v(\{i_k\})$ and rejects it otherwise.

At stage 5, as the representative of $N\backslash S$, i_n offers $z^{i_n}_{i_s} = v(S)$ to i_s, while i_s accepts any offer greater than or equal to $v(S)$ and rejects any offer strictly less than it.

Clearly the combination of these strategies of all players in N leads to acceptance at stage 3, which yields the Shapley value for any player who is not the proposer. Moreover, given that following the strategies the offer is accepted by all players and the grand coalition is formed, the proposer also obtains her Shapley value.

With the reasoning for stages 4 and 5 provided immediately before the proof, one can readily verify the above strategies are indeed best responses by superadditivity. To show that the strategies of the active players at stages 1, 2 and 3 are on a subgame perfect equilibrium path, we proceed by induction. The induction hypothesis is that whenever the game reaches a round with s active players comprising the set S, then in any SPE of the game induced by Mechanism A, each player $j \in S$ obtains $\phi_j(S, v|_S)$. To see this induction assumption holds for $s = 1$, note that when $s = 1$, the 'offer' (that is made by i_1 to himself) is vacuously accepted. Thus, the payoff to i_1 is $v(\{i_1\})$, which is indeed $\phi_{i_1}(\{i_1\}, v|_{\{i_1\}})$. We now assume it holds for any set of s active players, where $1 \leq s \leq n-1$, and then show that it holds for any set of $s+1$ active players.

To show that for an arbitrary set of $s+1$ active players, say $S \cup \{i_{s+1}\}$, each player $j \in S \cup \{i_{s+1}\}$ obtains $\phi_j(S \cup \{i_{s+1}\}, v|_{S\cup\{i_{s+1}\}})$ in SPE, please note the following: the strategies of active players at stage 3 are on an SPE path by the induction assumption; the strategy of the proposer at stage 2 is on an SPE path for a proposer that wishes to make an acceptable offer. More specifically, the reason that any proposer, other than the first rejected proposer i_n, would like to make an acceptable offer is superadditivity. Consequently, by backward induction, i_n can foresee that he can never obtain a higher payoff

ACCEPTED MANUSCRIPT

than $v(N) - v(N\backslash\{i_n\})$ in SPE. Hence, making an acceptable offer by the first proposer is part of an SPE. Finally, to check that the actions at stage 1, i.e., the bids, complete an SPE, note that all net bids equal zero by the balanced contributions property for the Shapley value (Myerson (1980)). Any deviation in the bids made by a player i cannot increase that player's payoff: lowering her total bids makes another player become the proposer, this will not change her payoff as all other players still bid the same way; raising her total bids can only lower her final payoff; maintaining the same level of her total bids does not improve her payoff.

The proof that any SPE yields the Shapley value involves induction as well. Although the induction hypothesis is essentially the same as the above one, we slightly modify it for the convenience of presentation. Here, the induction assumption is that whenever the game reaches a round with s active players comprising the set S, then in any SPE of the game induced by Mechanism A, each player $j \in S$ obtains $\phi_j(S, v|_S)$. This induction assumption apparently holds for $s = 1$. We now assume it holds for any set of $s - 1$ active players, where $2 \leq s \leq n$, and then show that it holds for any set of s active players by a series of claims. Please note that since this proof proceeds along similar lines as the unicity proof of Theorem 1 in Pérez-Castrillo and Wettstein (2001), we will omit its details. But for completeness, below we present these claims and provide some key explanations.

Claim (1). In any SPE at stage 5, i_n offers $z_{i_s}^{i_n} = v(S)$ to i_s, and i_s accepts any renegotiation offer $z_{i_s}^{i_n} \geq v(S)$ but rejects any $z_{i_s}^{i_n} < v(S)$.

Claim (2). In any SPE at stage 4, i_n offers $y_{i_k}^{i_n} = v(\{i_k\})$ to any subsequently rejected proposer i_k where $k = s+1, ..., n-1$. And any i_k, where $k = s+1, ..., n-1$, accepts any offer $y_{i_k}^{i_n} \geq v(\{i_k\})$ and rejects it otherwise.

Claim (3). In any SPE at stage 3, any player $j \in S\backslash\{i_s\}$ accepts any offer which is greater than or equal to $\phi_j(S\backslash\{i_s\}, v|_{S\backslash\{i_s\}})$ and rejects any offer strictly less than $\phi_j(S\backslash\{i_s\}, v|_{S\backslash\{i_s\}})$.

One can readily see that claims (1) and (2) follow directly from superadditivity, while claim (3) is due to the induction assumption.

Claim (4). In any SPE at stage 2, a proposer i_s, where $2 \leq s \leq n-1$, offers $x_j^{i_s} = \phi_j(S\backslash\{i_s\}, v|_{S\backslash\{i_s\}})$ to every $j \in S\backslash\{i_s\}$; and i_n may have two different strategies: one is to offer $x_j^{i_n} = \phi_j(N\backslash\{i_n\}, v|_{N\backslash\{i_n\}})$ to every $j \in N\backslash\{i_n\}$, and the other is to offer

ACCEPTED MANUSCRIPT

$x_j^{i_n} < \phi_j(N\backslash\{i_n\}, v|_{N\backslash\{i_n\}})$ to some $j \in N\backslash\{i_n\}$.

It is easy to see this is the best response for any subsequent proposer after i_n due to superadditivity, because any lower offer would be rejected, which results in the stand-alone payoff to him by claim (2). However, for the first proposer, i_n, it is still part of an SPE strategy to make an unacceptable offer at stage 2, so long as later at the renegotiation stage, he offers $z_{i_{n-1}}^{i_n} = v(N\backslash\{i_n\})$, which will result in the same payoff to i_n as in the case where he makes an acceptable offer at stage 2.

Claim (5). In any SPE, $B^i = B^j$ for all $i, j \in N$ and hence $B^i = 0$ for all $i \in N$.

Claim (6). In any SPE, each player's payoff is the same regardless of who is chosen as the proposer.

Claim (7). In any SPE, the final payoff received by each player coincides with each player's Shapley value.

The proofs of claims (5), (6) and (7) can be constructed in the same way as in Pérez-Castrillo and Wettstein (2001). □

Please note that in the above mechanism, even if the strict superadditivity is imposed, the SPE strategies are not unique, which is different from that of Pérez-Castrillo and Wettstein (2001). The flexibility of multiple SPE comes from the option of renegotiation. Contrasting with the SPE involving immediate acceptance by every $j \in N\backslash\{i_n\}$ at stage 3, a different SPE allows i_n to make an unacceptable offer (e.g., $x_j^{i_n} < \phi_j(N\backslash\{i_n\}, v|_{N\backslash\{i_n\}})$ to some $j \in N\backslash\{i_n\}$) at stage 2 but later, at the renegotiation stage, to offer $z_{i_{n-1}}^{i_n} = v(N\backslash\{i_n\})$, which restores the Shapley value outcome for every player.

We have seen that Mechanism A specifies a rule that at stage 5 it is the first rejected proposer, i_n, now as the representative of $N\backslash S$, who has the power to make an offer to i_s in renegotiation. In equilibrium obviously i_n will offer $v(S)$ to i_s, which makes stages 4 and 5 strategically redundant in generating the Shapley value beyond stage 3. Is this rule really necessary to obtain the Shapley value if we adopt such a renegotiation design? Alternatively, will a rule having the accepted proposer, i_s, now as the representative of S, make an offer to i_n in renegotiation lead to a different outcome from the Shapley value? Surprisingly, the Shapley value is robust to such a change-in-power in the bargaining with renegotiation.

ACCEPTED MANUSCRIPT

Mechanism B. The mechanism is identical structure-wise to Mechanism A. Stages $1, 2, 3$ and 4 are in effect the same as in Mechanism A. The description below is restricted to stage 5, where the difference with Mechanism A lies.

Stage 5: Player i_s makes a renegotiation offer $z_{i_n}^{i_s} \in \mathbb{R}$ to i_n who represents $N \backslash S$.

If the renegotiation offer is rejected, then the game ends with both sides, S and $N \backslash S$, staying apart. Thus, the payoff to each player is still the same as specified in the corresponding part of Mechanism A.

If the renegotiation offer is accepted, then i_s pays $z_{i_n}^{i_s}$ to i_n so that the grand coalition N is formed. Hence, the game ends and i_s finally receives $v(N) - z_{i_n}^{i_s} - \sum_{j \in S \backslash \{i_s\}} b_j^{i_s} - \sum_{j \in S \backslash \{i_s\}} x_j^{i_s} + \sum_{k=s+1}^{n} b_{i_s}^{i_k}$, each player $j \in S \backslash \{i_s\}$ finally receives $\sum_{k=s}^{n} b_j^{i_k} + x_j^{i_s}$, while i_n finally receives $z_{i_n}^{i_s} - z_{i_s}^{i_n} - \sum_{j \in N \backslash \{i_n\}} b_j^{i_n} - \sum_{k=s+1}^{n-1} y_{i_k}^{i_n}$ and each player i_k for $k = s+1, ..., n-1$ finally receives $y_{i_k}^{i_n} - \sum_{j \in N \backslash \{i_k, i_{k+1}, ..., i_n\}} b_j^{i_k} + \sum_{l \in \{i_{k+1}, ..., i_n\}} b_{i_k}^{l}$.

Theorem 3.2 *For any superadditive game (N, v), the SPE outcomes of Mechanism B coincide with the payoff vector $\phi(N, v)$ as prescribed by the Shapley value.*

Proof

The proof relies on the self-duality satisfied by the Shapley value. Due to superadditivity, at stage 5, the best response for i_s is to offer $z_{i_n}^{i_s} = v(N \backslash S)$ to i_n. Hence, in equilibrium, the amount of payoff that players in S bargain for equals $v(N) - v(N \backslash S)$ instead of $v(S)$ that was the case in Mechanism A. The proof thus proceeds as in that of Theorem 3.1, with the underlying characteristic function, describing the payoff to coalition S, given by $v^d(S) = v(N) - v(N \backslash S)$ (rather than $v(S)$). We know that (N, v^d) is the dual game of (N, v) and the Shapley value is self-dual, i.e., $\phi(v) = \phi(v^d)$. □

Comparing the above two renegotiation mechanisms, one may consider a compromised version (call the compromised mechanism) such that i_s and i_n bid for the right to make the renegotiation offer at stage 5. This will endogenize the selection of a proposer in renegotiation. Given Theorem 3.1 and Theorem 3.2, we immediately obtain the following result.

Corollary 3.3 *For any superadditive game (N, v), the SPE outcomes of the compromised mechanism coincide with the payoff vector $\phi(N, v)$ as prescribed by the Shapley value.*

The renegotiation mechanisms constructed so far require all the rejected proposers to wait until an active set of players S form into a coalition. Only by then will i_n be able to make a take-it-or-leave-it offer to all other rejected players to form a coalition $N \backslash S$ and

ACCEPTED MANUSCRIPT

can the game proceed to the renegotiation stage between S and $N\backslash S$. Now we consider a protocol such that the rejected proposers can form coalitions among themselves timely with no need to wait up to the moment S is formed. That is, soon after a proposer is rejected, he can make an offer to the proposer got rejected immediately before him. This mechanism seems even more realistic and practical especially when concerning the short time span and distance between the adjacent rejections and the neighboring rejected proposers. It is interesting that again the Shapley value emerges as the only SPE outcome of this mechanism.

Mechanism C. Stages 1 and 2 are the same as in Mechanism A. The description is, therefore, only focused on those differences appearing at stages 3, 4 and 5.

Stage 3: The players in S other than i_s, sequentially, either accept or reject the offer.

If the offer made at stage 3 is rejected, all players in S other than i_s proceed again through the mechanism from stage 1. Meanwhile, the rejected proposer i_s continues playing the game at stage 4.

If the offer of i_s is accepted, we have two cases where $S = N$ and $S \neq N$. If $S = N$, the game ends like in Mechanism A. If $S \neq N$, the game moves to stage 5a.

Stage 4: The rejected proposer i_s makes an offer $y^{i_s}_{i_{s+1}} \in \mathbb{R}$ to the previously rejected proposer i_{s+1}.

If the proposal is rejected by i_{s+1}, then i_s enters stage 5b by himself, whereas the final payoff to any player i_k where $k = s+1, ..., n$ is confirmed because he will not engage in any further negotiation. Note that i_k may have already formed with other rejected proposers into a coalition denoted by T^{i_k} (if i_k stands alone, $T^{i_k} = \{i_k\}$), then this player i_k finally receives $v(T^{i_k}) - y^{i_k}_{i_{k+1}} - \sum_{j\in N\backslash\{i_k,i_{k+1},...,i_n\}} b^{i_k}_j + \sum_{l\in\{i_{k+1},...,i_n\}} b^l_{i_k}$.

If the proposal is accepted by i_{s+1}, then i_s pays the offer $y^{i_s}_{i_{s+1}} \in \mathbb{R}$ to i_{s+1}. Note that if before accepting the offer from i_s, i_{s+1} has already formed a coalition with other rejected proposers, say $T^{i_{s+1}}$, then a larger coalition $T^{i_{s+1}} \cup \{i_s\}$ is established and i_s becomes its representative. The game proceeds to stage 5b.

Stage 5a: Player i_s, as a representative of S, makes a renegotiation offer $z^{i_s}_{i_{s+1}} \in \mathbb{R}$ to i_{s+1}. Note that i_{s+1} is the representative of coalition $T^{i_{s+1}} \subseteq N\backslash S$, although $T^{i_{s+1}}$ could contain i_{s+1} solely if i_{s+1} failed to have his offer be accepted by i_{s+2}.

If i_{s+1} rejects $z^{i_s}_{i_{s+1}}$, then the game ends with both sides, S and $T^{i_{s+1}}$, staying apart.

If i_{s+1} accepts $z^{i_s}_{i_{s+1}}$, then i_s pays $z^{i_s}_{i_{s+1}}$ to i_{s+1} so that the coalition $S \cup T^{i_{s+1}}$ is formed. Hence, the game ends and i_s finally receives $v(S \cup T^{i_{s+1}}) - z^{i_s}_{i_{s+1}} - \sum_{j\in S\backslash\{i_s\}} b^{i_s}_j - \sum_{j\in S\backslash\{i_s\}} x^{i_s}_j + \sum_{k=s+1}^{n} b^{i_k}_{i_s}$ and i_{s+1} finally receives $z^{i_s}_{i_{s+1}} - z^{i_{s+1}}_{i_{s+2}} - \sum_{j\in N\backslash\{i_{s+1},i_{s+2},...,i_n\}} b^{i_{s+1}}_j + \sum_{l\in\{i_{s+2},...,i_n\}} b^l_{i_{s+1}}$.

ACCEPTED MANUSCRIPT

Stage 5b: Player i_{s-1} makes an offer $\bar{z}_{i_s}^{i_{s-1}} \in \mathbb{R}$ to i_s. Here, dependent upon the result at stage 4, i_s can be an individual player or a representative of $T^{i_{s+1}} \cup \{i_s\}$.

Theorem 3.4 *For any superadditive game (N, v), the SPE outcomes of Mechanism C coincide with the payoff vector $\phi(N, v)$ as prescribed by the Shapley value.*

Proof

Duality again plays a key role here and the proof proceeds in a similar manner of Theorem 3.2. Hence, we only present the crucial part of the proof below.

Due to superadditivity, it is obvious that in equilibrium all rejected proposers will form into a coalition. That is, any rejected proposer i_k will offer $y_{i_{k+1}}^{i_k} = v(T^{i_{k+1}})$ to i_{k+1}, and consequently, the coalition is gradually expanded until $N\backslash S$ is formed before i_s makes a renegotiation offer to i_{s+1} who is the representative of $N\backslash S$. Therefore, at stage 5a, the best response for i_s is to offer $z_{i_{s+1}}^{i_s} = v(N\backslash S)$ to i_{s+1}. Consequently, in equilibrium, the amount of payoff that players in S are actually bargaining for equals $v(N) - v(N\backslash S)$.

Thus, if an offer made by i_s can be accepted by players in $S\backslash\{i_s\}$ at stage 3, it necessarily requires the sum to be no less than $v(N) - v((N\backslash S) \cup \{i_s\})$. Next, by offering $z_{i_{s+1}}^{i_s} = v(N\backslash S)$ to i_{s+1} as specified above, i_s receives $v(N) - (v(N) - v((N\backslash S) \cup \{i_s\})) - v(N\backslash S) = v((N\backslash S) \cup \{i_s\}) - v(N\backslash S)$ at stage 5a. On the other hand, if i_s was rejected at stage 3, he will become the representative of $(N\backslash S) \cup \{i_s\}$ at stage 4 in equilibrium (due to superadditivity), and (also) receives $v((N\backslash S) \cup \{i_s\}) - v(N\backslash S)$ at stage 5b.

Comparing the above two cases, we can see that i_s is indifferent between making an offer be accepted or rejected by S at stage 3. Consequently, we know that all SPE will be in the following format, i.e., it does not matter whether a proposer i_s makes an offer be accepted or rejected at stage 3.

At stage 1, each player $i \in S$ announces $b_j^i = \phi_j(S, v^d|_S) - \phi_j(S\backslash\{i\}, v^d|_{S\backslash\{i\}})$ for every $j \in S\backslash\{i\}$, where $(S, v^d|_S)$ is defined by $v^d|_S(T) = v(N) - v(N\backslash T)$ for all $T \subseteq S$.

At stage 2, a proposer, player i_s, can make any offer $x_j^{i_s} \in \mathbb{R}$ to every $j \in S\backslash\{i_s\}$.

At stage 3, any player $j \in S\backslash\{i_s\}$ accepts any offer which is greater than or equal to $\phi_j(S\backslash\{i_s\}, v^d|_{S\backslash\{i_s\}})$ and rejects any offer strictly less than $\phi_j(S\backslash\{i_s\}, v^d|_{S\backslash\{i_s\}})$.

At stage 4, i_s makes an offer $y_{i_{s+1}}^{i_s} = v(T^{i_{s+1}})$ to i_{s+1}. And i_{s+1} accepts an offer $y_{i_{s+1}}^{i_s} \geq v(T^{i_{s+1}})$ and rejects it otherwise.

At stage 5a, as the representative of S, i_s offers $z_{i_{s+1}}^{i_s} = v(T^{i_{s+1}})$ to i_{s+1}, while i_{s+1} accepts any offer greater than or equal to $v(T^{i_{s+1}})$ and rejects it otherwise.

ACCEPTED MANUSCRIPT

At stage 5b, as the representative of T^{i_s}, i_s accepts any offer $\bar{z}_{i_s}^{i_{s-1}} \geq v(T^{i_s})$ from i_{s-1} and rejects it otherwise, while i_{s-1} offers $\bar{z}_{i_s}^{i_{s-1}} = v(T^{i_s})$.

Thus, the final payoff to i_s, independent of the acceptance or rejection of his offer at stage 3, equals

$$\begin{aligned}
& v((N\backslash S)\cup\{i_s\}) - v(N\backslash S) - \sum_{j\in S\backslash\{i_s\}} b_j^{i_s} + \sum_{k=s+1}^{n} b_{i_s}^{i_k} \\
= \; & v((N\backslash S)\cup\{i_s\}) - v(N\backslash S) - \sum_{j\in S\backslash\{i_s\}} \big(\phi_j(S, v^d|_S) - \phi_j(S\backslash\{i\}, v^d|_{S\backslash\{i\}})\big) \\
& + \sum_{k=s+1}^{n} \big(\phi_{i_s}(N\backslash\{i_{k+1},...,i_n\}, v^d|_{N\backslash\{i_{k+1},...,i_n\}}) - \phi_{i_s}(N\backslash\{i_k,...,i_n\}, v^d|_{N\backslash\{i_k,...,i_n\}})\big) \\
= \; & v((N\backslash S)\cup\{i_s\}) - v(N\backslash S) - \big(v(N) - v(N\backslash S) - \phi_{i_s}(S, v^d|_S)\big) \\
& + (v(N) - v((N\backslash S)\cup\{i_s\})) + \big(\phi_{i_s}(N, v^d) - \phi_{i_s}(N\backslash\{i_{s+1},...,i_n\}, v^d|_{N\backslash\{i_{s+1},...,i_n\}})\big) \\
= \; & \phi_{i_s}(N, v^d) = \phi_{i_s}(N, v).
\end{aligned}$$

For those who are not chosen as a proposer, it is obvious to see that their final payoffs are equal to the Shapley value of (N, v). □

Based on the SPE strategy profile provided above, we like to highlight an intriguing feature of this mechanism: The Shapley value can be realized as an SPE outcome in the very beginning of the game (where at stage 2 the first proposer i_n offers $x_j^{i_n} = \phi_j(N\backslash\{i_n\}, v^d|_{N\backslash\{i_n\}})$ to every $j \neq i_n$), or at any time later (e.g., after all players $i_{s+1}, ..., i_n$ have been rejected, i_s makes an offer be rejected at stage 3 and further acts according to the strategy specified above). Hence, the mechanism allows players to go through the entire equilibrium path without necessarily stopping at the beginning of the game, while the Shapley value can be retained. This further shows that, even with the condition of strict superadditivity, unlike in Pérez-Castrillo and Wettstein (2001), the SPE strategies of the above mechanisms are not unique despite their outcomes are the same. The Shapley value can be realized in an SPE where no renegotiation is actually played, or can be reached in an SPE where a proposer made his offer be rejected but later 'successfully' played at the renegotiation stage.

As one can see, the only technical requirement we adopted to obtain the above results is the strict superadditivity on TU environment. One may wonder whether we can relax this assumption to an arbitrary TU game, like the way adopted in Pérez-Castrillo and Wettstein (2001) by using the Shapley value of the superadditive cover of the game. The answer is no because the superadditive cover does not necessarily rule out equality relations between the values of coalitions, which might lead to unacceptable offers looking for higher

ACCEPTED MANUSCRIPT

payoffs from the renegotiation stage rather than the formation of the grand coalition. On the other hand, if a TU game is such that its superadditive cover satisfies the strict superadditivity, then all the above results will hold.

4 Concluding remarks

The paper aims to develop new insights in strategic coalitional bargaining, and to push forward this research with a fresh treatment in modeling, where players are allowed to further negotiate and form coalitions after they were rejected in the first instance. This more realistic setup shows a certain robustness of the Shapley value in non-cooperative bargaining. One more original aspect lies in the use of the self-duality of the Shapley value, which plays a key role in proving the main results of the paper. This follows the spirit of the Nash program as we can better understand an axiomatic property by discovering its relevant strategic features. Moreover, as we can see from the above mechanisms (see especially the comment after Theorem 3.4), allowing options for rejected players to further negotiate and form coalitions is not only interesting in restoring the Shapley value per se, but also helps to yield the Shapley value in a much more flexible manner: many more SPE can emerge and they do not require immediate acceptance at the beginning of the game.

Regarding the renegotiation protocol, one may naturally ask a question: Why not simply let the rejected proposers play again the bidding stage and offer stage, like the incumbent players do? The reason is that such a construction may lead to endless repetition without a proper stopping rule of the game. One can of course deal with this issue by introducing a discount factor or breakdown probability, which will end up in a similar situation to the one studied in Hart and Mas-Colell (1996). Hence, without modeling with a discount factor, the rejected players should organize themselves in a different bargaining rule from that for the incumbent players so that the entire bargaining can stop properly. And this does not seem far from reality. After all, two organizations rarely use identical rules.

Finally, for future research, the renegotiation idea of the paper may be useful for one to construct non-cooperative bargaining protocols for other self-dual solution concepts.

References

Gul, F. (1989) "Bargaining Foundations of Shapely Value", *Econometrica*, 57, 81-95.

Gul, F. (1999) "Efficiency and immediate agreement: a reply to Hart and Levy", *Econometrica*, 67, 913-918.

ACCEPTED MANUSCRIPT

Hart, S., and A. Mas-Colell , (1996), "Bargaining and Value", *Econometrica*, 64, 357-380.

Jackson, M.O. (2001) "A crash course in implementation theory", *Social Choice and Welfare*, 18, 655-708.

Ju, Y., and D. Wettstein (2009) "Implementing Cooperative Solution Concepts: a Generalized Bidding Approach", *Economic Theory*, 39, 307-330.

Kalai, E. and D. Samet (1987) "On weighted Shapley values", *International Journal of Game Theory*, 16, 205-222.

Krishna, V. and R. Serrano (1995) "Perfect equilibria of a model of n-person non-cooperative bargaining", *International Journal of Game Theory*, 27, 259-272.

Maskin, E. (2003), "Bargaining, coalitions and externalities", *Presidential Address of the Econometric Society*.

Nash, J.F. (1950a), "Equilibrium points in n-person games", *Proceedings of the National Academy of Sciences*, 36, 48-49.

Nash, J.F. (1950b), "The bargaining problem", *Econometrica*, 18, 155-162.

Nash, J.F. (1953), "Two person cooperative games", *Econometrica*, 21, 128-140.

Pérez-Castrillo, D. and D. Wettstein (2001) "Bidding for the Surplus: a Non-Cooperative Approach to the Shapley Value", *Journal of Economic Theory*, 100, 274-294.

Pérez-Castrillo, D. and D. Wettstein (2002) "Choosing wisely: a multi-bidding approach", *American Economic Review*, 92, 1577-1587.

Serrano, R. (2005), "Fifty years of the Nash program, 1953-2003", *Investigaciones Economicas*, 29, 219-258.

Shapley, L.S. (1953), "A value for n-person games", in: Kuhn, H.W., Tucker, A.W. (eds.) *Contributions to the Theory of Games II*, Princeton University Press, Princeton, NJ, pp. 307-317.

Trockel, W. (2002), "Integrating the Nash Program into Mechanism Theory", *Review of Economic Design*, 7, 27-43.

Sun, N., W. Trockel and Z. Yang (2008) "Competitive Outcomes and Endogenous Coalition Formation in an n-Person Game", *Journal of Mathematical Economics*, 44, 853-860.