Top responsiveness and stable partitions in coalition formation games^{*}

Dinko Dimitrov[†]

Institute of Mathematical Economics Bielefeld University Germany

Shao Chin Sung Department of Industrial and Systems Engineering

Aoyama Gakuin University

Japan

March 2005

Abstract

Top responsiveness is introduced by Alcalde and Revilla [Journal of Mathematical Economics 40 (2004) 869-887] as a property which induces a rich domain on players's preferences in hedonic games, and guarantees the existence of core stable partitions. We strengthen this observation by proving the existence of strict core stable partitions, and when a mutuality condition is imposed as well, the existence of Nash stable partitions.

JEL classification: C71; C72; C78; D71

Keywords: Coalition formation; Nash stability; Partitions; Strict core stability

^{*}The authors gratefully acknowledge financial support from CentER, Tilburg University, where the work on this paper has started. The work of D. Dimitrov was also supported by a Humboldt Research Fellowship conducted at Bielefeld University.

[†]Corresponding author. E-mail address: d.dimitrov@wiwi.uni-bielefeld.de

1 Introduction

One possibility to study the process of coalition formation is to model it as a hedonic coalition formation game (cf. Banerjee et al. (2001) and Bogomolnaia and Jackson (2002)). In such a model each player's preferences over coalitions depend solely on the composition of members of her coalition (cf. Drèze and Greenberg (1980)). Given a hedonic game, the main interest is then in the existence of outcomes (partitions of the set of players) that are stable in some sense. For example, the focus in the work of Banerjee et al. (2001) is on the existence of core stable partitions, while Bogomolnaia and Jackson (2002) present sufficient conditions for the existence of Nash and individually stable partitions.

Top responsiveness is introduced by Alcalde and Revilla (2004) as a condition on players' preferences, which captures the idea of how each player believes that others could complement her in the formation of research teams. As shown by these authors, top responsiveness is a sufficient condition for the existence of core stable partitions in hedonic games. This result is provided constructively, i.e., an algorithm, called the *top covering algorithm*, is proposed for generating a core stable partition. As argued by Alcalde and Revilla (2004), top responsiveness induces a rich domain on players' preferences and economic agents' behavior seems to be consistent with this property. Therefore it is interesting to study, on the induced preference domain, the existence of partitions that are stable also with respect to other stability notions like *strict core stability*, Nash stability, and *individual stability*.

In doing so, we strengthen the results of Alcalde and Revilla (2004) in two ways. First, we show the existence of strict core stable partitions, and hence, the existence of individually stable partitions is implied. Second, we consider an example to illustrate that a Nash stable partition may fail to exist, and prove that imposing a mutuality condition turns out to be sufficient for the existence of Nash stable partitions. Indeed, we show that the partition generated by a simplified version of the top covering algorithm is strict core stable and when the mutuality condition is imposed the partition is Nash stable as well, where the simplified version of the top covering algorithm always returns the same outcome as the top covering algorithm. Since the top covering algorithm is shown by Alcalde and Revilla (2004) to be strategyproof on the induced preference domain, the results in this paper can be seen as completing the study of hedonic games under top responsiveness with respect to stability and manipulability issues.

2 Basic setup

2.1 Hedonic games

Let $N = \{1, 2, ..., n\}$ be a finite set of players. Each nonempty subset of N is called a *coalition*. Each player i is endowed with preferences \succeq_i over the set $\mathcal{A}^i = \{X \subseteq N \mid i \in X\}$ of all possible coalitions she may belong to, i.e., each \succeq_i is a complete pre-ordering over \mathcal{A}^i . A *hedonic game* is described by a pair $\langle N, \succeq \rangle$, where \succeq is a profile of players' preferences, i.e., $\succeq = (\succeq_1, \succeq_2, \ldots, \succeq_n)$. An *outcome* Π for $\langle N, \succeq \rangle$ is a partition of the player set N, i.e., Π is a collection of nonempty pairwise disjoint coalitions whose union is N. For each partition Π of N and for each player $i \in N$, we denote by $\Pi(i)$ the coalition in Π containing i.

Given a hedonic game $\langle N, \succeq \rangle$ and a partition Π of N, we say that¹

¹ Notice that these stability notions are defined in a "positive" way that will be very

• Π is *core stable* if, for each nonempty $X \subseteq N$,

 $-\Pi(i) \succeq_i X$ for some $i \in X$;

• Π is strictly core stable if, for each nonempty $X \subseteq N$,

 $-\Pi(i) \succeq_i X$ for each $i \in X$ if $X \succeq_i \Pi(i)$ for each $i \in X$;

• Π is Nash stable if, for each $X \in \Pi \cup \{\emptyset\}$ and for each $i \in N$,

 $- \Pi(i) \succeq_i X \cup \{i\};$

• Π is *individually stable* if, for each $X \in \Pi \cup \{\emptyset\}$ and for each $i \in N$,

 $- X \succ_j X \cup \{i\}$ for some $j \in X$ if $X \cup \{i\} \succ_i \Pi(i)$.

Observe that strict core stability implies core stability, and that Nash stability implies individual stability. Moreover, it can easily be verified that strict core stability implies individual stability as well.

2.2 Choice sets and top responsiveness

Let $i \in N$ and $X \in \mathcal{A}^i$. We denote by $Ch(i, X) \subseteq 2^X \cap \mathcal{A}^i$ the set of maximals of i on X under \succeq_i , i.e.,

$$Ch(i, X) = \{ Y \in 2^X \cap \mathcal{A}^i \mid Y \succeq_i Z \text{ for each } Z \in 2^X \cap \mathcal{A}^i \}.$$

Observe that each $Y \in Ch(i, X)$ satisfies $i \in Y \subseteq X$. Moreover, for each $Y, Z \subseteq 2^X \cap \mathcal{A}^i$, we have $Y \succ_i Z$ if $Y \in Ch(i, X)$ and $Z \notin Ch(i, X)$.

As in the work of Alcalde and Revilla (2004), we assume that players' preferences satisfy top responsiveness, i.e., we assume that, for each $i \in N$ the following three conditions are satisfied:

useful when providing our existence proofs in the next section.

Condition 1: For each $X \in \mathcal{A}^i$, |Ch(i, X)| = 1.

By ch(i, X) we denote the unique maximal set of player i on X under \succeq_i , i.e., $Ch(i, X) = \{ch(i, X)\}$. Then,

- **Condition 2:** For each pair $X, Y \in \mathcal{A}^i, X \succ_i Y$ if $ch(i, X) \succ_i ch(i, Y)$;
- **Condition 3:** For each pair $X, Y \in \mathcal{A}^i, X \succ_i Y$ if $ch(i, X) \sim_i ch(i, Y)$ and $X \subset Y$.

Suppose Condition 1 is fulfilled, and suppose $ch(i, X) \sim_i ch(i, Y)$ and $X \subset Y$. Then, we have $ch(i, X) \in Ch(i, Y)$. From Condition 1, we have $Ch(i, Y) = \{ch(i, Y)\}$, and thus, ch(i, X) = ch(i, Y). Hence, Condition 3 can be reformulated as follows.

• for each pair $X, Y \in \mathcal{A}^i, X \succ_i Y$ if ch(i, X) = ch(i, Y) and $X \subset Y$.

2.3 The simplified top covering algorithm

In order to show that top responsiveness is a sufficient condition for the existence of a core stable partition, Alcalde and Revilla (2004) propose an algorithm, called the *top covering algorithm*, which can be seen as a generalization of Gale's top trading cycle (see Shapley and Scarf (1974) for more details). In the following, a simplified version of this algorithm is described.

Let t be a positive integer. We define a function $C^t : N \times 2^N \to 2^N$ as follows. For each $i \in N$ and for each $X \in \mathcal{A}^i$,

- $\mathcal{C}^1(i, X) = ch(i, X)$, and
- $\mathcal{C}^{t+1}(i, X) = \bigcup_{j \in \mathcal{C}^t(i, X)} ch(j, X)$ for each positive integer t.

Let $i \in N$ and $X \in \mathcal{A}^i$. Observe that $j \in ch(j, X) \subseteq X$ if $j \in X$ (i.e., $X \in \mathcal{A}^j$), and thus, $i \in \mathcal{C}^1(i, X) \subseteq X$. Let t be a positive integer, and suppose $\mathcal{C}^t(i, X) \subseteq X$. Then, $j \in ch(j, X) \subseteq X$ for each $j \in \mathcal{C}^t(i, X)$, and by definition, $\mathcal{C}^t(i, X) \subseteq \mathcal{C}^{t+1}(i, X) \subseteq X$. It follows that $\mathcal{C}^{|N|+1}(i, X) =$ $\mathcal{C}^{|N|}(i, X)$. By $\mathcal{CC}(i, X)$ we denote $\mathcal{C}^{|N|}(i, X)$. Then, we have

$$i \in ch(i, X) \subseteq \mathcal{CC}(i, X) \subseteq X$$
 for each $i \in X$. (1)

Moreover, from $\mathcal{C}^{|N|+1}(i, X) = \mathcal{C}^{|N|}(i, X) = \mathcal{CC}(i, X)$, one can easily show by induction on t that $\mathcal{C}^t(j, X) \subseteq \mathcal{CC}(i, X)$ for each $j \in \mathcal{CC}(i, X)$ and for each positive integer t. It follows that

$$\mathcal{CC}(j,X) \subseteq \mathcal{CC}(i,X)$$
 for each $j \in \mathcal{CC}(i,X)$. (2)

Now we are ready to describe the simplified top covering algorithm.

Simplified top covering algorithm:

Given: A hedonic game $\langle N, \succeq \rangle$ satisfying top responsiveness.

Step 1: Set $R^1 := N$ and $\Pi := \emptyset$.

Step 2: For k := 1 to |N|:

Step 2.1: Select an i ∈ R^k satisfying |CC(i, R^k)| ≤ |CC(j, R^k)| for each j ∈ R^k.
Step 2.2: Set S^k := CC(i, R^k), Π := Π ∪ {S^k}, and R^{k+1} := R^k \ S^k.
Step 2.3: If R^{k+1} = Ø, then goto Step 3.

Step 3: Return Π as outcome.

We denote by $\Pi^{TC}_{\langle N, \succeq \rangle}$ the outcome obtained by applying the simplified top covering algorithm to $\langle N, \succeq \rangle$.

Lemma 1 Let $\langle N, \succeq \rangle$ be a hedonic game satisfying top responsiveness. When applied to $\langle N, \succeq \rangle$, the simplified top covering algorithm ends in finite steps and its outcome $\Pi_{\langle N, \succeq \rangle}^{TC}$ is a partition of N.

Proof. It is obvious that the simplified top covering algorithm ends in finite steps, because Step 2 repeats at most |N| times.

To show that $\Pi_{\langle N, \succeq \rangle}^{TC}$ is a partition of N, suppose Step 2 repeats $1 \leq K \leq |N|$ times. Then, we have $\Pi_{\langle N, \succeq \rangle}^{TC} = \{S^1, S^2, \ldots, S^K\}$ and $S^1 \cup S^2 \cup \cdots \cup S^K \cup R^{K+1} = N$. From Step 2.2, the coalitions $S^1, S^2, \ldots, S^K, R^{K+1}$ are pairwise disjoint, and moreover from (1), the coalitions S^1, S^2, \ldots, S^K are all nonempty. Hence, $\Pi_{\langle N, \succ \rangle}^{TC}$ is a partition of N if $R^{K+1} = \emptyset$.

Indeed, when K < |N|, we have $R^{K+1} = \emptyset$ from Step 2.3. When K = |N|, we have $|S^1 \cup S^2 \cup \cdots \cup S^K| \ge K = |N|$ from the disjointedness and nonemptiness of S^k s, which implies $R^{K+1} = \emptyset$.

In order to analyze the partition $\Pi_{\langle N, \succeq \rangle}^{TC}$, for each $i \in N$, we denote by k(i)the number such that $i \in S^{k(i)}$. In other words, the coalition $\Pi_{\langle N, \succeq \rangle}^{TC}(i) = S^{k(i)}$ is included into $\Pi_{\langle N, \succeq \rangle}^{TC}$ at the k(i)th iteration of Step 2. Since $\Pi_{\langle N, \succeq \rangle}^{TC}$ is a partition of N, the number k(i) is well-defined for each $i \in N$.

Lemma 2 Let $\langle N, \succeq \rangle$ be a hedonic game satisfying top responsiveness. Then, for each $i \in N$, $\Pi^{TC}_{\langle N, \succeq \rangle}(i) = CC(i, R^{k(i)})$.

Proof. Let $i \in N$, and let $j \in R^{k(i)}$ be the player selected at the k(i)th iteration of Step 2.1. Then, we have k(j) = k(i) and $i \in \Pi_{\langle N, \succeq \rangle}^{TC}(i) = \mathcal{CC}(j, R^{k(j)}) = \mathcal{CC}(j, R^{k(i)})$, and thus, from Step 2.1 we have $|\mathcal{CC}(j, R^{k(i)})| \leq |\mathcal{CC}(i, R^{k(i)})|$. On the other hand, we have $\mathcal{CC}(i, R^{k(i)}) \subseteq \mathcal{CC}(j, R^{k(i)})$ from $i \in \mathcal{CC}(j, R^{k(i)})$ and (2). Therefore, we have $\mathcal{CC}(i, R^{k(i)}) = \mathcal{CC}(j, R^{k(i)})$, which

implies $\Pi^{TC}_{\langle N, \succeq \rangle}(i) = \mathcal{CC}(i, R^{k(i)}).$

From Lemma 2 and (1), we have

$$ch(i, R^{k(i)}) \subseteq \Pi^{TC}_{\langle N, \succeq \rangle}(i) \quad \text{for each } i \in N.$$
 (3)

3 Results

3.1 Strict core stability

For a hedonic game $\langle N, \succeq \rangle$ satisfying top responsiveness, our first result relates a partition obtained by the simplified top covering algorithm with the set of strictly core stable partitions of the player set N. Since strict core stability implies individual stability, it follows by Theorem 1 stated below that the simplified top covering algorithm generates an individually stable partition as well.

Theorem 1 Let $\langle N, \succeq \rangle$ be a hedonic game satisfying top responsiveness. Then, $\Pi_{\langle N, \succeq \rangle}^{TC}$ is strictly core stable for $\langle N, \succeq \rangle$.

Proof. Suppose there exists a nonempty $X \subseteq N$ such that $X \succeq_i \Pi_{\langle N, \succeq \rangle}^{TC}(i)$ for each $i \in X$, and we show that $\Pi_{\langle N, \succeq \rangle}^{TC}(i) \succeq_i X$ for each $i \in X$.

In doing so, we will apply Condition 3 based on the fact that for $j \in X$ with $k(j) \leq k(i)$ for each $i \in X$ the following three statements hold:

$$X \succeq_{j} \Pi^{TC}_{\langle N, \succeq \rangle}(j), \tag{\alpha}$$

$$ch(j,X) = ch(j,\Pi^{TC}_{\langle N,\succeq \rangle}(j)), \qquad (\beta)$$

$$\Pi^{TC}_{\langle N,\succeq\rangle}(j) \subseteq X. \tag{(\gamma)}$$

We have to show (β) and (γ) since (α) is given by supposition.

To show (β) , notice that from $k(j) \leq k(i)$ for each $i \in X$ it follows that $X \subseteq R^{k(j)}$. Hence, from (3), we have $ch(j, R^{k(j)}) \subseteq \Pi^{TC}_{\langle N, \succeq \rangle}(j) = \mathcal{CC}(j, R^{k(j)}) \subseteq R^{k(j)}$, which implies $ch(j, \Pi^{TC}_{\langle N, \succeq \rangle}(j)) = ch(j, R^{k(j)})$. By (α) and Condition 2 we have $ch(j, X) \succeq_j ch(j, \Pi^{TC}_{\langle N, \succeq \rangle}(j))$. Hence, from $X \subseteq R^{k(j)}$, we have

$$ch(j,X) = ch(j,\Pi_{\langle N,\succeq\rangle}^{TC}(j)) = ch(j,R^{k(j)}), \qquad (4)$$

i.e., (β) holds true.

We turn now to showing (γ) . Since, by Lemma 2, $\Pi_{\langle N, \succeq \rangle}^{TC}(j) = \mathcal{CC}(j, \mathbb{R}^{k(j)})$, we can show (γ) by showing that $\mathcal{C}^t(j, \mathbb{R}^{k(j)}) \subseteq X$ for each positive integer t. Let t = 1. Then, from (4), we have

$$\mathcal{C}^1(j, R^{k(j)}) = ch(j, R^{k(j)}) = ch(j, X) \subseteq X.$$

Suppose $C^t(j, R^{k(j)}) \subseteq X$ for some positive integer t. Observe that $\Pi^{TC}_{\langle N, \succeq \rangle}(i) = \Pi^{TC}_{\langle N, \succeq \rangle}(j) = CC(j, R^{k(j)})$ for each $i \in C^t(j, R^{k(j)})$. From (1) and (2), $ch(i, R^{k(j)}) \subseteq CC(i, R^{k(j)}) \subseteq CC(j, R^{k(j)}) = \Pi^{TC}_{\langle N, \succeq \rangle}(i)$ for each $i \in C^t(j, R^{k(j)})$, and thus, $ch(i, \Pi^{TC}_{\langle N, \succeq \rangle}(i)) = ch(i, R^{k(j)})$. Again, by assumption, for each $i \in C^t(j, R^{k(j)})$, we have $X \succeq_i \Pi^{TC}_{\langle N, \succeq \rangle}(i)$, and we have $ch(i, X) \succeq_i ch(i, \Pi^{TC}_{\langle N, \succeq \rangle}(i))$ from Condition 2. Hence, from $X \subseteq R^{k(j)}$, we have $ch(i, X) = ch(i, \Pi^{TC}_{\langle N, \succeq \rangle}(i)) = ch(i, R^{k(j)})$. It follows that $ch(i, R^{k(j)}) \subseteq X$ for each $i \in C^t(j, R^{k(j)})$, and moreover,

$$\mathcal{C}^{t+1}(j, R^{k(j)}) = \bigcup_{i \in \mathcal{C}^t(j, R^{k(j)})} ch(i, R^{k(j)}) \subseteq X.$$

Therefore, $\mathcal{C}^t(j, \mathbb{R}^{k(j)}) \subseteq X$ for each positive integer t, and thus,

$$\Pi^{TC}_{\langle N,\succeq \rangle}(j) = \mathcal{CC}(j, R^{k(j)}) \subseteq X,$$

i.e., (γ) holds true.

Finally, notice that from (α) , (β) , and (γ) we have $X = \prod_{\langle N, \succeq \rangle}^{TC}(j)$ by Condition 3. Therefore, $\prod_{\langle N, \succeq \rangle}^{TC}(i) = \prod_{\langle N, \succeq \rangle}^{TC}(j) \succeq_i X$ for each $i \in X$.

3.2 Nash stability

Nash stability is a stronger stability notion than individual stability since in its definition there are no requirements on a positive reaction of the welcoming coalition. Unfortunately, top responsiveness does not guarantee the existence of a Nash stable partition as exemplified next.

Example 1 Let $N = \{1, 2, 3\}$ and players' preferences be as follows:

{1}	\succ_1	$\{1, 2\}$	\sim_1	$\{1, 3\}$	\succ_1	$\{1, 2, 3\},\$
$\{2\}$	\succ_2	$\{1,2\}$	\sim_2	$\{2, 3\}$	\succ_2	$\{1, 2, 3\},\$
$\{1, 2, 3\}$	\succ_3	$\{1, 3\}$	\sim_3	$\{2, 3\}$	\succ_3	$\{3\}$.

The reader can easily check that this game satisfies top responsiveness. Notice that any partition in which player 1 and player 2 are not single will be blocked by the corresponding player. Hence, we have to check only the partition $\Pi = \{\{1\}, \{2\}, \{3\}\}\}$. However, $\{1,3\} \succ_3 \{3\}$ (and $\{2,3\} \succ_3 \{3\}$), i.e. a Nash stable partition does not exist for this game.

In order to guarantee the existence of a Nash stable partition for a hedonic game $\langle N, \succeq \rangle$ we will require, in addition to top responsiveness, $\langle N, \succeq \rangle$ to satisfy *mutuality*, i.e., the following condition:

• For each $i, j \in N$ and for each $X \in \mathcal{A}^i \cap \mathcal{A}^j$, $i \in ch(j, X)$ if and only if $j \in ch(i, X)$.

In other words, mutuality requires that, for any group X of players, the members of every player's maximal on X mutually complement each other. In the formulation of this condition we were inspired by the existence result in the seminal paper of Bogomolnaia and Jackson (2002). These authors show that the combination of additive separability and symmetry (a stronger version of mutuality) guarantees the existence of Nash stable partitions. Here, we show that, by imposing both top responsiveness and mutuality, the simplified top covering algorithm generates a Nash stable partition. Note that one can easily construct a hedonic game in which players' preferences satisfy top responsiveness but are not additive separable.

Theorem 2 Let $\langle N, \succeq \rangle$ be a hedonic game satisfying top responsiveness and mutuality. Then, $\Pi^{TC}_{\langle N, \succeq \rangle}$ is Nash stable for $\langle N, \succeq \rangle$.

Proof. Suppose $\Pi_{\langle N, \succeq \rangle}^{TC}$ is not Nash stable for $\langle N, \succeq \rangle$, i.e., there exist $X \in \Pi_{\langle N, \succeq \rangle}^{TC} \cup \{\emptyset\}$ and $i \in N$ such that $X \cup \{i\} \succ_i \Pi_{\langle N, \succeq \rangle}^{TC}(i)$. Since $\Pi_{\langle N, \succeq \rangle}^{TC}$ is strictly core stable (by Theorem 1), we have

$$\Pi^{TC}_{\langle N,\succeq\rangle}(i) \succeq_i \{i\},\$$

which implies $X \neq \emptyset$. From $X \cup \{i\} \succ_i \Pi^{TC}_{\langle N, \succeq \rangle}(i)$, we have $X \neq \Pi^{TC}_{\langle N, \succeq \rangle}(i)$, which implies $X \cap \Pi^{TC}_{\langle N, \succeq \rangle}(i) = \emptyset$ and thus $i \notin X$.

Observe that $X \not\subseteq R^{k(i)}$. Suppose otherwise, i.e., $X \subseteq R^{k(i)}$. From $X \cup \{i\} \succ_i \Pi^{TC}_{\langle N, \succeq \rangle}(i)$ and $ch(i, R^{k(i)}) \subseteq CC(i, R^{k(i)}) = \Pi^{TC}_{\langle N, \succeq \rangle}(i)$, we have $ch(i, X \cup \{i\}) = ch(i, R^{k(i)})$. Since $X \cap \Pi^{TC}_{\langle N, \succeq \rangle}(i) = \emptyset$, we have $ch(i, X \cup \{i\}) = \{i\}$, and from Condition 3 and $X \neq \emptyset$, we have $\Pi^{TC}_{\langle N, \succeq \rangle}(i) \succeq_i \{i\} \succ_i X \cup \{i\}$, which contradicts to our supposition.

Let k be a positive integer such that $S^k = X$. Then, we have $X \subseteq R^k$ and k(j) = k for each $j \in X$, and from Lemma 2, $\mathcal{CC}(j, R^k) = X$ for each $j \in X$. Moreover, from (1), we have $ch(j, R^k) \subseteq \mathcal{CC}(j, R^k) = X$ for each $j \in X$, and from $i \notin X$, we have $i \notin ch(j, R^k)$ for each $j \in X$. From $X \notin R^{k(i)}$, we have k < k(i), which implies $i \in R^{k(i)} \subseteq R^k$. Then, we have $ch(j, R^k) = ch(j, X \cup \{i\}) = ch(j, X)$, and thus,

$$i \notin ch(j, X \cup \{i\})$$
 for each $j \in X$.

By mutuality, we have $j \notin ch(i, X \cup \{i\})$ for each $j \in X$, which implies $ch(i, X \cup \{i\}) = \{i\}$. Again, from Condition 3 and $X \neq \emptyset$, we have $\Pi^{TC}_{\langle N, \succeq \rangle}(i) \succeq_i \{i\} \succ_i X \cup \{i\}, \text{ which contradicts again to our supposition.} \blacksquare$

References

- Alcalde, J. and P. Revilla (2004): Researching with whom? Stability and manipulation, Journal of Mathematical Economics 40, 869-887.
- [2] Banerjee, S., H. Konishi, and T. Sönmez (2001): Core in a simple coalition formation game, Social Choice and Welfare 18, 135-153.
- [3] Bogomolnaia, A. and M.O. Jackson (2002): The stability of hedonic coalition structures, Games and Economic Behavior 38, 201-230.
- [4] Dréze, J. and J. Greenberg (1980): Hedonic coalitions: optimality and stability, Econometrica 48, 987-1003.
- [5] Shapley, L. and H. Scarf (1974): On cores and indivisibility, Journal of Mathematical Economics 1, 23-28.