TOPS RESPONSIVENESS, STRATEGY-PROOFNESS AND COALITION FORMATION PROBLEMS*

José Alcalde and Pablo Revilla**

WP-AD 2001-11

Correspondence to J. Alcalde: Universidad de Alicante, Departamento de Fundamentos de Análisis Económico, Campus San Vicente del Raspeig, s/n, 03071 Alicante (Spain). Tel.: +34 96 590 3400 x 3236 / Fax: +34 96 590 3898 / E-mail: alcalde@merlin.fae.ua.es.

Editor: Instituto Valenciano de Investigaciones Económicas, S.A.
First Edition April 2001.
Depósito Legal: V-2042-2001

---

* The authors are grateful to Anna Bogomolnaia, Katarina Cechlávorá, Guillaume Haeringer, Jordi Massó, Andrés Perea, Antonio Romero-Medina and Guadalupe Valera for their comments. This work is partially supported by DGCYT project PB 97-0131, and the Institut Valencià d'Investigacions Econòmiques.

** J. Alcalde: University of Alicante; P. Revilla: University Pablo Olavide (Sevilla).

TOPS RESPONSIVENESS, STRATEGY-PROOFNESS AND COALITION FORMATION PROBLEMS

José Alcalde and Pablo Revilla

A B S T R A C T

This paper introduces a property over agents' preferences, called Tops Responsiveness Condition. Such a property guarantees that the core in Hedonic Coalition Formation games is not empty. It is also shown that a mechanism exists that selects a stable allocation. It turns out that this mechanism, to be called tops covering, is strategy-proof even if the core is not a singleton. Furthermore, we also find out that the tops covering mechanism is the only strategy-proof mechanism that always selects stable allocations.

KEYWORDS: Coalition Formation; Stability; Strategy Proofness.

1. Introduction

Economic agents usually cooperate to reach some objectives that they could not obtain by themselves. It seems to be clear that agents' cooperation is funded on the own interest of each economic agent. We might give a further step saying that such behavior is followed by agents not only by economic purposes but also for other (more generic) reasons.

In real-live situations one can find several examples that can be (partially) explained by the above principle. For instance, economic researchers could increase the quality level of their findings by their cooperation on researching. Workers form associations, called unions, to defend collectively their interests faced to firms' desires on work contractual conditions. The emergence of the European Economic Union can be explained on the basis of countries' interest on reaching some economic and/or political agreements that involve their cooperation on certain aspects. Similar arguments can be used for the cases of the North-American agreements signed with the name of NAFTA and, most recently, MERCOSUR.

The above examples present some similarities which allow us to give a common treatment to all of them. In fact we can say that all these examples can be expressed in generic terms as follows. A set of agents, who could find profitable to reach agreements on cooperating, decide to form coalitions under which this fruitful cooperation will be done.

The study of how some of these agreements emerge can be done on the basis of a cooperative game-theoretical approach. There is ample literature on the study of how coalitions are formed. Some authors focussed on the case in which the players only have preferences regarding the members of their coalition but not, however, regarding its entire structure. The agents involved in a coalition tend to behave as though they were immune to the decisions made by other coalitions. This phenomenon, that Drèze and Greenberg (1980) called the hedonic aspect of coalition formation, is the sort of game we focus on in this paper.

We deal with a static study that can be viewed as an initial approach to understanding why some coalitions are formed. We shall focus specifically on two main aspects of the implementation of such agreements, namely, their stability and the existence of sincere mechanisms implementing them.

We consider an agreement to be stable whenever it is impossible to find a set of agents that could block it, i.e., they cannot find another agreement, involving only these agents, that is advantageous for all of them. Unfortunately, in a general setting, it is impossible to guarantee the existence of stable agreements. This problem was pointed out by Gale and Shapley (1962) showing that the set of stable agreement configurations may be empty, whenever only two-agents agreements could be signed.

Some authors concentrate on the study of economic environments in which stable configurations always exist. In such a context, Alcalde (1995) solves the existence of stable configurations for two-agents agreements. The concept of the above-mentioned paper is generalized by Banerjee et al. (2001), who consider economic environments that cannot be described as a Cartesian product of the agents' characteristics. This description of their environments imposes additional difficulties on the study of comparative statics (see Alcalde and Romero-Medina (2000) for a discussion on this subject). Cechlárová and Romero-Medina (1998) provide an economic environment to guarantee the existence of stable agreement configurations. The description of the agents' preferences in such a case is quite simple. Each individual orders the set of agents, without including itself. When an individual has to compare two sets of agents with whom she could reach an agreement, she cares only about the best partners, according to the above ordering. Recently, Alcalde and Romero-Medina (2000) have described economic environments, based on individual preferences, under which stable agreement configurations always exist.

In this paper we deal with the possibility of generalizing the idea beyond the concept of Essentiality, defined by Alcalde and Romero-Medina (2000). This will be done with the Tops Responsiveness condition, a property that guarantees the existence of stable allocations, whenever the agents' preferences fulfill it.

The second question to be dealt with, is the analysis of the agents' strategic behavior when they have to follow some specific rule in order to decide which agreements they should sign. Following the framework proposed by Gibbard (1973) and Sattherwaite (1975), we concentrate on direct mechanisms, i.e., agents are asked to declare their own preferences on the set of outcomes (agreements configurations), and a fixed rule will select an outcome for any preferences profile. We intend to avoid conclusions that would recommend employing rules that are difficult to justify. With such a view, we shall require that our rules fulfill two basic properties: stability and strategy-proofness. By strategy-proofness we ensure that no agent could benefit from lying about her own preferences when an agreement rule is employed. Formal definitions for both of these concepts will be introduced later on.

The main results reported in the literature regarding the problem of finding stable rules that would avoid any strategic behavior by the agents involved are rather pessimistic. In fact, Gibbard (1973) and Sattherwaite (1975) state that the two properties are incompatible. In *matching models*, a framework which is quite close to the one we present here, Roth (1982) obtains negative results which are strengthened by Alcalde and Barberà (1994).

Such results, however, can not be applied directly to our model, as they were defined for mechanisms that are applied to universal domains, i.e. no restriction is imposed on the agents' preferences. Note that, in our framework, an impossibility result can be straightforwardly found since, for some cases, there is no stable allocation. In our model, however, when the Tops Responsiveness Condition is fulfilled by the agents' preferences, we design a mechanism that associates a stable agreement configuration to each economic environment and, as it turns out, the mechanism we provide is strategy-proof.

The paper is organized as follows: Section 2 outlines the basic framework and introduces the Tops Responsiveness Condition, a property that describes the economic environments that will be analyzed throughout the paper. Section 3 shows the existence of stable agreement configurations under the Tops Responsiveness Condition. Section 4 introduces a mechanism, to be called the *Tops Covering Mechanism*, which selects a stable agreement configuration for each economic environment that satisfies the Tops Responsiveness Condition. We also show that such a mechanism satisfies strategy-proofness. It turns out that stability and strategy-proofness, together, characterize the Tops Covering Mechanism. Conclusions are gathered to Section 5. Finally, an appendix contains some technical proofs.

2. The Framework

Let $N = \{1, \ldots, i, \ldots, n\}$ be the set of agents. A subset S of N is called an agreement. Let $\mathcal{A}^i = \{S \subseteq N : i \in S\}$ be the set of agreements that agent i can reach. Each agent i is endowed with a complete pre-ordering R_i, defined over $\mathcal{A}^i$, which represents her preferences over all the possible agreements she can reach. We denote, by P_i the strict preference derived from R_i, i.e. SP_iS' means that SR_iS' and not $S'R_iS$. An agreements problem will be shortly described by a pair $\{N, R\}$, where the agents' preferences profile R is a description of each agent's preferences.

A solution for an agreements problem, also called an *agreements configuration*, is a function $\mu : N \mapsto 2^N$ such that

(*i*) $\mu(i) \in \mathcal{A}^i$, for each $i \in N$, and (*ii*) $\mu(i) \cap \mu(j) \neq \emptyset$ if, and only if, $\mu(i) = \mu(j)$ for any i, $j \in N$.

In fact, an agreements configuration can be viewed as a partition of the set of agents.

We say that an agreements configuration μ is *stable* for $\{N, R\}$ if there is no non-empty set of agents, say S, such that $SP_i\mu(i)$ for each $i \in S$. A set satisfying the above property is said to *block* agreement μ. Finally, we state that an agreements configuration μ is *individually rational* for $\{N, R\}$ if there is no single agent blocking it.

In this paper we focus on the study of stable agreements configurations for any given problem $\{N, R\}$. Since we are not concerned with any analysis of population variations, we shall consider N to be fixed throughout the paper.

An agreements configuration rule γ is a function that selects, for each possible preferences profile R, an agreements configuration for $\{N, R\}$. We say that rule γ is stable if it selects a stable agreements configuration for each problem, i.e. for any preferences profile R, $\gamma(R)$ is stable for $\{N, R\}$.

We will provide an example showing the general impossibility of finding stable rules. In this example, we build a situation that has no stable allocation.

Example 2.1. *A problem with no stable agreement configuration.*
Let $N = \{1, 2, 3\}$ *with the preferences described in the following table:*

$$\begin{array}{ccccccc} \{1,2\} & P_1 & \{1,3\} & P_1 & \{1\} & P_1 & \{1,2,3\} \\ \{2,3\} & P_2 & \{1,2\} & P_2 & \{2\} & P_2 & \{1,2,3\} \\ \{1,3\} & P_3 & \{2,3\} & P_3 & \{3\} & P_3 & \{1,2,3\} \end{array}$$

This problem has no stable allocation. Note that the "autarchic" situation, in which $\mu(i) = \{i\}$ *for each* i*, is blocked by any pair of agents. Agreement* μ' *in which* $\mu'(i) = \{1, 2, 3\}$ *for each agent* i *is blocked by any agent. Finally, any agreement* μ'' *where* $\mu''(i) = \{i\}$ *for some* i *and* $\mu''(j) = \mu''(k) = \{j, k\}$ *for the rest is blocked by a coalition formed by* i *and another agent. If* $i = 1$, $\{1, 3\}$ *is the blocking pair;* $\{1, 2\}$ *will be the blocking pair if* $i = 2$*; and* $\{2, 3\}$ *blocks* μ'' *when* $i = 3$. ■

As we mentioned in the Introduction, the aim of some recent papers has been the study of economic environments in which stable allocations always exist. We introduce the "Tops Responsiveness Condition," a property that is weaker than both Essentiality and the Union Responsiveness Condition (Alcalde and Romero-Medina, 2000). This property may or may not be satisfied by any agent's preferences. But, as Theorem 3.4 informs us, when each agent has preferences satisfying *Tops Responsiveness*, the set of stable agreements configurations is non-empty.

Definition 2.2. *We say that agent i's preferences, R_i, satisfy Tops Responsiveness on $\mathcal{A}^i$, if for any set S in $\mathcal{A}^i$ there is only a maximal for R_i, to be denoted by $Ch_i(S)$;*[1] *and for any two sets S and S' in $\mathcal{A}^i$, the following conditions are fulfilled:*

1. *$Ch_i(S)\,P_i Ch_i(S')$ implies SP_iS', and*
2. *If $Ch_i(S) = Ch_i(S')$, and $S \subset S'$ then SP_iS'.*[2]

Let $\mathcal{R}$ denote the set of agents' preference profiles, where each agent's preferences satisfy the Tops Responsiveness Condition.

Let us illustrate the two conditions referred to above with the following example: We shall consider five researchers, A, B, C, D, and E, who may or may not be able to reach any agreements on cooperating on writing a paper. Agent A thinks that her cooperation with B and C will produce the nicest paper she can do. Condition 1 means that A will also prefer any agreement involving all of the five researchers (or any four containing A, B and C) to any other that is not agreed to by B (or C). The idea is that agent A needs the cooperation of B and C to produce such a nice paper, and this is what does matter to such an agent! Condition 2 states that agent A would prefer to have an agreement with researchers B, C and D, rather than the one involving all of the other researchers. This is because A thinks that (a) incorporating more researchers will not improve the quality of the paper, and (b) the more authors a paper have, the lower importance is attributed to each one. To sum-up, therefore, it is more important for A to avoid the absence of B or C than to prevent the presence of any other researcher, (D or E). In other words, the idea beyond the Tops Responsiveness condition can be illustrated by the sentence "my friends' friends are my friends," and considering that each one just prefers to stay with her own friends, so the least friends' friends come to the group, the better I feel.

[1] $Ch_i(S)$ denotes the choice of agent i in set S under preferences R_i. Thus, $Ch_i(S)$ is the set $S' \subseteq S$ such that $S'P_iS''$ for any other set $S'' \subseteq S$.

[2] In order to avoid confusions, let us note that, throughout this paper, we use the symbol $\subset$ for the strict inclusion (i.e. $A \subset B$ does not allow the case where $A = B$), whereas the symbol $\subseteq$ will be used for the weak inclusion.

3. Tops Responsiveness and Stability

We have stated in Section 2 that any agreements problem whose agents' preferences satisfy the Tops Responsiveness Condition have stable agreements configurations. The aim of this section is to provide a formal statement for this fact. We will introduce a procedure which selects a stable agreements configuration for any problem whose agents' preferences exhibit the Tops Responsiveness property. This procedure can be viewed as a natural extension for the $\mathcal{RA}$-*algorithm* by Alcalde and Romero-Medina (2000) (and henceforth to the Gale's Tops Trading Cycle introduced in Shapley and Scarf (1974)).

Before introducing our rule, to be called the *tops covering algorithm*, we need some additional notation. Let us consider an agreements problem $\{N, R\}$ satisfying the Tops Responsiveness Condition. Given two sets of agents, S and S', with $S \subseteq S'$, let define $Ch_S(S') = \cup_{i \in S} Ch_i(S')$. For each agent i, and set S in $\mathcal{A}^i$, let $\tau_i(S)$ be the smaller set containing $Ch_i(S)$ such that $Ch_{\tau_i(S)}(S) = \tau_i(S)$.[3] Note that such a set can be built straightforwardly by using the following *choices covering algorithm*, whose output is $\tau_i(S)$.

Definition 3.1. *The Choices Covering Algorithm works as follows: Let $S \subseteq N$ be a set of agents, and i be an agent in S. Let us assume that all agents in S exhibit preferences satisfying the Tops Responsiveness Condition.*

Step 1. *Let $S^1 = Ch_i(S)$. If $Ch_{S^1}(S) = S^1$, then let $\tau_i(S) = S^1$, otherwise, go to Step 2.*

Step k. *Let $S^k = Ch_{S^{k-1}}(S)$. If $Ch_{S^k}(S) = S^k$, then let $\tau_i(S) = S^k$, otherwise, go to Step k+1.*

Note that for any $k \geq 2$, $S^{k-1} \subseteq S^k \subseteq S$. Since N contains finite elements, this property guarantees that the above algorithm always converges in a finite number of steps.

Definition 3.2. *Let $\{N, R\}$ be an agreements problem whose agents' preferences satisfy the Tops Responsiveness Condition. The Tops Covering Algorithm works as follows:*

[3] Henceforth, $\tau_i(S)$ has to fulfill two requirements: (a) $Ch_{\tau_i(S)}(S) = \tau_i(S)$, and (b) $Ch_T(S) \neq T$ for any $T \subset \tau_i(S) \cap \mathcal{A}^i$.

Step (1) *Let compute, for each agent* i*, the set* $\tau_i(N)$*. For each agent* i *such that* $\tau_i(N) \subseteq \tau_h(N)$ *for each* $h \in \tau_i(N)$*, let set* $\mu_i^1 = \tau_i(N)$*. Finally, let* $N^1 = \{i \in N : \mu_i^1 = \tau_i(N)\}$*. If* $N^1 = N$*, stop; otherwise, go to Step (2).*

Step (2) *Let compute, for each agent* $i \in N \backslash N^1$*, the set* $\tau_i(N \backslash N^1)$*. For each agent* i *such that* $\tau_i(N \backslash N^1) \subseteq \tau_h(N \backslash N^1)$ *for each* $h \in \tau_i(N \backslash N^1)$*, let set* $\mu_i^2 = \tau_i(N \backslash N^1)$*. Finally, let* $N^2 = \{i \in N \backslash N^1 : \mu_i^2 = \tau_i(N \backslash N^1)\}$*. If* $N^2 = N \backslash N^1$*, stop; otherwise, go to Step (3).*

Step (k) *Let compute, for each agent* $i \in N \backslash \cup_{j<k} N^j$*, the set* $\tau_i(N \backslash \cup_{j<k} N^j)$*. For each agent* i *such that* $\tau_i(N \backslash \cup_{j<k} N^j) \subseteq \tau_h(N \backslash \cup_{j<k} N^j)$ *for each* $h \in \tau_i(N \backslash \cup_{j<k} N^j)$*, let set* $\mu_i^k = \tau_i(N \backslash \cup_{j<k} N^j)$*. Finally, let* $N^k = \left\{i \in N \backslash \cup_{j<k} N^j : \mu_i^k = \tau_i(N \backslash \cup_{j<k} N^j)\right\}$*. If* $N = \cup_{j \leq k} N^j$*, stop; otherwise, go to Step (k+1).*

The algorithm ends at the Step t-th satisfying that $\cup_{j \leq t} N^j = N$*, and produce an output* $\mu^{tc}(R)$ *such that, for each set* N^j *and agent* i *in* N^j*,* $\mu_i^{tc}(R) = \mu_i^j$*.*

In order to show how the previous algorithm works, let us consider the following four agents' problem.

Example 3.3. *Let* $N = \{1, 2, 3, 4\}$*, with agents' preferences*

$$\begin{array}{l} \{1,2\} P_1 \{1,2,3\} P_1 \{1,2,4\} P_1 \{1,2,3,4\} P_1 \{1,3\} P_1 \{1,3,4\} P_1 \{1\} P_1 \{1,4\} \\ \{1,2,3\} P_2 \{1,2,3,4\} P_2 \{1,2\} P_2 \{1,2,4\} P_2 \{2\} P_2 \{2,4\} P_2 \{2,3\} P_2 \{2,3,4\} \\ \{1,3,4\} P_3 \{1,2,3,4\} P_3 \{2,3,4\} P_3 \{3\} P_3 \{2,3\} P_3 \{1,3\} P_3 \{1,2,3\} P_3 \{3,4\} \\ \{4\} P_4 \{3,4\} P_4 \{2,4\} P_4 \{1,4\} P_4 \{1,2,4\} P_4 \{1,3,4\} P_4 \{2,3,4\} P_4 \{1,2,3,4\} \end{array}$$

In this case, we have that $\tau_1(N) = \tau_2(N) = \tau_3(N) = N$*, and* $\tau_4(N) = \{4\}$*, hence* $N^1 = \{4\}$*, and* $\mu_4^{tc}(R) = \mu_4^1 = \{4\}$*. In the second Step, we have that* $\tau_1(\{1,2,3\}) = \tau_2(\{1,2,3\}) = \{1,2,3\}$*, and* $\tau_3(\{1,2,3\}) = \{3\}$*, so* $N^2 = \{3\}$*, and* $\mu_3^{tc}(R) = \mu_3^2 = \{3\}$*. Finally, at the third Step,* $\tau_1(\{1,2\}) = \tau_2(\{1,2\}) = \{1,2\}$*, which yields* $N^3 = \{1,2\}$*, and the output* $\mu_1^{tc}(R) = \mu_1^3 = \mu_2^{tc}(R) = \mu_2^3 = \{1,2\}$*. As can be seen, the output of our procedure is a stable agreement configuration for the original problem* $\{N, R\}$*.*

A nice feature of the Tops Covering Algorithm is that it produces a stable agreement for any problem whose agents' preferences satisfy the Tops Responsiveness Condition. The conclusion of the next result is a direct consequence of this fact.

Theorem 3.4. *Let $\{N, R\}$ be an agreements problem whose agents' preferences satisfy the Tops Responsiveness Condition, then the set of stable agreement configurations is non-empty.*

Proof. See the appendix. ■

4. Strategic Behavior and the Tops Covering Algorithm

This section is devoted to the analysis of agents' behavior when they are forced to follow some specific rule in order to establish whether an agreement can be signed. We insist on two main requirements to consider a rule acceptable: stability and strategy-proofness. Stability means that the only rules that we shall consider are those that, for any given problem $\{N, R\}$, assign a stable agreement configuration, relative to R. Strategy-proofness means that the rule is immune to agents' manipulations.

Definition 4.1. *Let N be a fixed set of agents, and let $\mathcal{R}^*$ be a set of agents' preference profiles. A stable agreements mechanism on $\mathcal{R}^*$ is a function μ^* which selects, for each R in $\mathcal{R}^*$, an agreements configuration $\mu^*(R)$, which is stable for $\{N, R\}$.*

Definition 4.2. *Let N be a fixed set of agents, and let $\mathcal{R}^*$ be a set of agents' preference profiles. We say that mechanism μ^* is strategy-proof on $\mathcal{R}^*$ if for all R in $\mathcal{R}^*$, and each $R_i^{'}$ in $\mathcal{R}_i^*$ we have that*[4]

$$\mu_i^*(R)\, R_i \mu_i^* \left(R_{-i}, R_i^{'}\right)$$

In our study, given that we want the rule to always provide a stable allocation, we have to restrict attention to the frameworks where such an allocation exists. We have already seen that when agents' preferences satisfy the Tops Responsiveness Condition, stable agreements configurations exist. Hence, we can formulate this question: Let us consider that agents' preferences are forced to satisfy the Tops Responsiveness Condition. Is there any strategy-proof rule selecting stable agreements configurations? A positive answer to this question is given in Theorem 4.3.

[4] In what follows, we use the classical formulation. Given a preferences profile R, $\left(R_{-i}, R_i^{'}\right)$ will denote the profile where agent i's preferences, R_i, have been replaced by $R_i^{'}$.

Theorem 4.3. *Let N be a fixed set of agents. Let define the Tops Covering Mechanism as the mechanism which selects, for each R in $\mathcal{R}$ the agreements configuration $\mu^{tc}(R)$. Then, the Tops Covering Mechanism is strategy-proof.*

Proof. Let us assume that for some agreement problem, say $\{N, R\}$, there is an agent $\hat{\imath}$ and preferences for this agent $R_{\hat{\imath}}^{'}$ such that $\mu_{\hat{\imath}}^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right) P_{\hat{\imath}} \mu_{\hat{\imath}}^{tc}(R)$.

For each agent i, let $k(i)$ denote the stage of the algorithm (when applied to preferences R) in which this agent reaches her agreement.[5] Let us note that, for any agent i such that $k(i) < k(\hat{\imath})$, we have that the set τ_i computed for the society $N \backslash \cup_{1 \leq k < k(i)} N^k$, does not depends on agent $\hat{\imath}$'s preferences:

$$\tau_i\left(N \backslash \cup_{1 \leq k < k(i)} N^k; R\right) = \tau_i\left(N \backslash \cup_{1 \leq k < k(i)} N^k; R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right).$$

This implies that, for each i such that $k(i) < k(\hat{\imath})$, $\mu_i^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right) = \mu_i^{tc}(R)$. Henceforth, $\mu_{\hat{\imath}}^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right) \subseteq N \backslash \cup_{1 \leq k < k(\hat{\imath})} N^k$. Moreover, $\forall i \in \mu_{\hat{\imath}}^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right) \backslash \{\hat{\imath}\}$ it is satisfied that $Ch_i\left(N \backslash \cup_{1 \leq k < k(\hat{\imath})} N^k\right) \subseteq \mu_i^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right)$. Since $\mu_{\hat{\imath}}^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right) P_{\hat{\imath}} \mu_{\hat{\imath}}^{tc}(R)$, the fact that agents $\hat{\imath}$'s preferences satisfy the Tops Responsiveness Condition implies that it must be the case that $Ch_{\hat{\imath}}\left(N \backslash \cup_{1 \leq k < k(\hat{\imath})} N^k\right) \subseteq \mu_i^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right)$. Therefore, by definitions 3.1 and 3.2, $\mu_{\hat{\imath}}^{tc}(R) \subseteq \mu_i^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right)$, which contradicts that $\mu_{\hat{\imath}}^{tc}\left(R_{-\hat{\imath}}, R_{\hat{\imath}}^{'}\right) P_{\hat{\imath}} \mu_{\hat{\imath}}^{tc}(R)$. ■

We now deal with the characterization of stable rules satisfying strategy-proofness in the family of problems whose agents' preferences satisfy the Tops Responsiveness Condition. An answer to that question can be trivially found in some environments. For instance, Alcalde and Romero-Medina (2000) show that, when agents' preferences satisfy Essentiality (a property stronger than the Tops Responsiveness Condition), there is a unique stable allocation. Henceforth, we can trivially state that, when agents' preferences are forced to satisfy essentiality, the Tops Covering Mechanism is the only stable and strategy-proof mechanism. The reason for this is quite simple: The mechanism satisfies both properties, and since the set of stable allocations is always a singleton, any stable mechanism would have to coincide with ours. This simple argument, however, can not be applied to the more general setting of agents' preferences satisfying the Tops Responsiveness Condition. This is because, as the following example shows, the set of stable agreement configurations is not generally a singleton.

[5] Following the notation employed in Definition 3.2, $k(i)$ is the integrer such that $i \in N^{k(i)}$.

Example 4.4. *Let $N = \{1, 2, 3\}$, with preferences*

$$\{1,2\}\, P_1 \{1,2,3\}\, P_1 \{1\}\, P_1 \{1,3\},$$
$$\{2,3\}\, P_2 \{1,2,3\}\, P_2 \{2\}\, P_2 \{1,2\},$$
$$\{1,3\}\, P_3 \{1,2,3\}\, P_3 \{2,3\}\, P_3 \{3\}$$

In this problem, there are two stable agreements configurations. $\mu^{tc}(R) = [\{1, 2, 3\}]$, and $\mu'(R) = [\{1\}, \{2, 3\}]$.

This example is also useful in demonstrating why any stable and strategy-proof mechanism must select the Tops Covering agreements configuration. Let us note that, in this example $\mu_3^{tc}(R)\, P_3 \mu_3^{'}(R)$. If we are employing any stable mechanism selecting, for this problem, $\mu'(R)$, agent 3 can manipulate by declaring preferences R_3' such that

$$\{1,3\}\, P_3 \{1,2,3\}\, P_3 \{3\}\, P_1 \{2,3\}$$

This fact can be extended to any stable mechanism selecting, for some problem, an agreements configuration different from one proposed by the Tops Covering mechanism. This is the aim of Theorem 4.5.

Theorem 4.5. *Let N be a fixed set of agents, and let Ψ be a stable mechanism on $\mathcal{R}$. If Ψ is strategy-proof, then $\Psi(R) = \mu^{tc}(R)$ for all R in $\mathcal{R}$.*

Before giving a formal proof for Theorem 4.5, let us introduce a way of describing how agents could manipulate. The idea beyond this possibility of manipulation is somehow similar to the one used in an impossibility result due to Alcalde and Barberà (1994).

Let R be a preferences profile in $\mathcal{R}$. Given any agent i, let R_i^{tc} denote agent i's preferences defined by:

1. For any two sets S and S' in $\mathcal{A}^i \setminus \{i\}$, $SR_i^{tc}S'$ if, and only if, SR_iS', and
2. For any set S in $\mathcal{A}^i \setminus \{i\}, \{i\}\, P_i^{tc} S$ if, and only if, $\mu_i^{tc}(R)\, P_i S$ and $\mu_i^{tc}(R) \nsubseteq S$.

Note that R_i^{tc} can be understood as the agent i's preferences obtained from R_i, by stating that any agreement configuration that i considers worse that $\mu_i^{tc}(R)$ is now to be considered individually irrational from i's point of view, the only exception comes from supersets of $\mu_i^{tc}(R)$, in order to guarantee that R_i^{tc} satisfies Tops Responsiveness.

It is straightforward to see that $\mu^{tc}(R) = \mu^{tc}(R')$ for any preferences profile R' where $R_i^{'} \in \{R_i, R_i^{tc}\}$.

Proof of Theorem 4.5 Let Ψ be a strategy-proof stable mechanism on $\mathcal{R}$, and let us assume that $\Psi \neq \mu^{tc}$. Hence, there should be a preferences profile R such that $\Psi(R) \neq \mu^{tc}(R)$.

Since both $\Psi(R)$ and $\mu^{tc}(R)$ are stable for R, and $\Psi(R) \neq \mu^{tc}(R)$, there should be an agent i preferring $\Psi_i(R)$ rather than $\mu_i^{tc}(R)$. Moreover, by the stability of the above agreements configurations, it should be the case that for each i in N, such that $\Psi_i(R)\ P_i\ \mu_i^{tc}(R)$, there should be an agent i' in $\Psi_i(R)$ such that $\mu_i^{tc}(R)\, P_{i'} \Psi_i(R)$.

The rest of the proof is done by an inductive argument on the agents' strategies. Since the arguments on how the agents can manipulate reach a high level of sophistication as the number of agents increases, we provide the formal arguments not only for the first agent, but for a few stages. This is done just to show how this induction is done.

Let N^k be the set of agents identified in Step (k) of the Tops Covering Algorithm when applied to $\{N, R\}$. (See Definition 3.2.) Without loss of generality, let us assume that there is an agent i belonging to N^1 such that $\Psi_i(R)\ P_i\ \mu_i^{tc}(R)$.[6] It should, therefore, be the case that $Ch_i(N) \subseteq \Psi_i(R) \subset \mu_i^{tc}(R)$. By construction of $\mu_i^{tc}(R)$, and given that μ^{tc} satisfies individual rationality, we have that $\mu_i^{tc}(R) \neq \{i\}$. Hence, there must be an agent, say i_1 such that

$$i_1 \in \mu_i^{tc}(R) \cap \Psi_i(R)\text{, and}$$

$$Ch_{i_1}(N) \nsubseteq \Psi_i(R)\text{,}$$

hence $\mu_i^{tc}(R)\, P_{i_1} \Psi_i(R)$.

Now, let us consider that i_1 states preferences $R_{i_1}^{tc}$. Note that, in such a case, any agreements configuration μ with $Ch_{i_1}(N) \nsubseteq \mu(i_1)$, and $\mu(i_1) \neq \{i_1\}$, fails to be stable for $\left\{N, \left(R_{-i_1}, R_{i_1}^{tc}\right)\right\}$. Hence, we have that

1. $\Psi_{i_1}\left(R_{-i_1}, R_{i_1}^{tc}\right) = \{i_1\}$, or
2. $Ch_{i_1}(N) \subseteq \Psi_{i_1}\left(R_{-i_1}, R_{i_1}^{tc}\right)$.

[6] Note that if there is not such an agent in N^1, we can proceed as follows. We select R such that there is an agent in N^1 for which $\Psi_i(R) \neq \mu_i^{tc}(R)$ holds. If such a profile does not exist, we select the preferences profile R for which the agent with the lowest $k(i)$ fulfills this property. Note that $\Psi_i(R) = \mu_i^{tc}(R)$ will hold for all agents in N^k for $k < k(i)$. Hence, the argument throughout the rest of the proof can be followed by considering $N \setminus \cup_{k<k(i)} N^k$ instead of N.

If $Ch_{i_1}(N) \subseteq \Psi_{i_1}\left(R_{-i_1}, R^{tc}_{i_1}\right)$, then i_1 could manipulate Ψ at R via $R^{tc}_{i_1}$. Note that, in such a case, the Tops Responsiveness Condition states that

$$\Psi_{i_1}\left(R_{-i_1}, R^{tc}_{i_1}\right) P_{i_1} \Psi_{i_1}(R).$$

Therefore, strategy-proofness satisfied by Ψ, implies that $\Psi_{i_1}\left(R_{-i_1}, R^{tc}_{i_1}\right)$ must be $\{i_1\}$.

Furthermore, given that μ^{tc}_i is not a singleton, there should be an agent i_2 in $\mu^{tc}_i(R)$ such that $i_1 \in Ch_{i_2}(N)$. This comes from the fact that $i_1 \in N^1$, and μ^{tc} satisfies individual rationality.

Given that $\mu^{tc}(R) = \mu^{tc}\left(R_{-i_1}, R^{tc}_{i_1}\right)$, we have that

$$\mu^{tc}_i\left(R_{-i_1}, R^{tc}_{i_1}\right) P_{i_2} \Psi_{i_2}\left(R_{-i_1}, R^{tc}_{i_1}\right).$$

Hence, by applying the same argument as we did above, we have that, if agent i_2 states preferences $R^{tc}_{i_2}$, it should be the case that $\Psi_{i_2}\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right) = \{i_2\}$.

At this stage, we must consider the following two possibilities:

(a_2) $\mu^{tc}_i(R) = \{i_1, i_2\}$. Then, $Ch_{i_1}(N) = Ch_{i_2}(N) = \{i_1, i_2\}$, which contradicts that $\Psi\left(R_{-i_1}, R^{tc}_{i_1}\right)$ was stable for $\left\{N, \left(R_{-i_1}, R^{tc}_{i_1}\right)\right\}$, or

(b_2) $\mu^{tc}_i(R) \neq \{i_1, i_2\}$. Then, there should be an agent, i_3 in $\mu^{tc}_i(R) \setminus \{i_1, i_2\}$ such that $\{i_1, i_2\} \cap Ch_{i_3}(N) \neq \emptyset$.

Let select agent i_3 such that $i_2 \in Ch_{i_3}(N)$.

Then, since $\Psi_{i_2}\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right) = \{i_2\}$, it must be the case that

$$Ch_{i_3}(N) \nsubseteq \Psi_{i_3}\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right).$$

If there is no agent i_3 such that $i_2 \in Ch_{i_3}(N)$, then choose agent i_3 such that $i_1 \in Ch_{i_3}(N)$.

Given that $i_2 \in \mu^{tc}_i(R)$, it must be the case that $i_2 \in Ch_{i_1}(N)$. Note that, in this case, we have

$$\Psi_{i_2}\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right) = \{i_2\}, \text{ and}$$

$$\{i_1\} P^{tc}_{i_1} S \text{ for any } S \in \mathcal{A}^{i_1} \text{ not containing } i_2,$$

hence

$$Ch_{i_3}(N) \nsubseteq \Psi_{i_3}\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right).$$

Moreover, $\mu^{tc}(R) = \mu^{tc}\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right)$. Let us now consider that agent i_3's preferences were $R^{tc}_{i_3}$. From the arguments provided above, it must be that

$\Psi_{i_3}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right) = \{i_3\}$. Notice that, otherwise, i_3 might manipulate Ψ at $\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right)$ via $R^{tc}_{i_3}$.

We should now consider, again, the following two cases:

(a_3) $\mu^{tc}_i(R) = \{i_1, i_2, i_3\}$. Then,

$$\cup^3_{k=1} Ch_{i_k}(N) = \{i_1, i_2, i_3\},$$

which contradicts that $\Psi\left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right)$ was stable for the problem $\left\{N, \left(R_{-\{i_1,i_2\}}, R^{tc}_{i_1}, R^{tc}_{i_2}\right)\right\}$.

(b_3) $\mu^{tc}_i(R) \neq \{i_1, i_2, i_3\}$. Then, there should be an agent, i_4 in $\mu^{tc}_i(R)\backslash\{i_1, i_2, i_3\}$ such that $\{i_1, i_2, i_3\} \cap Ch_{i_4}(N) \neq \emptyset$.

Let us select agent i_4 in accordance with the following priorities rule.

[1] $i_3 \in Ch_{i_4}(N)$. Then, since

$$\Psi_{i_3}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right) = \{i_3\},$$

we have that $Ch_{i_4}(N) \subsetneqq \Psi_{i_4}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right)$.

[2] If there is no agent satisfying [1], it must be that $i_3 \in Ch_{i_1}(N)$, or $i_3 \in Ch_{i_2}(N)$. We then select i_4 such that

[α] $i_2 \in Ch_{i_4}(N)$ if $i_3 \in Ch_{i_2}(N)$. Note that, then we have that $Ch_{i_4}(N) \subsetneqq \Psi_{i_4}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right)$ because

$$\Psi_{i_3}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right) = \{i_3\},$$

and

$$\{i_2\}\, P^{tc}_{i_2} S \text{ for any } S \in \mathcal{A}^{i_2} \text{ not containing } i_3,$$

or, if no agent satisfies [α],

[β] $i_1 \in Ch_{i_4}(N)$ if $i_3 \in Ch_{i_1}(N)$. Since

$$\Psi_{i_3}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right) = \{i_3\},$$

and

$$\{i_1\}\, P^{tc}_{i_1} S \text{ for any } S \in \mathcal{A}^{i_1} \text{ not containing } i_3,$$

we can also conclude, once more, that

$$Ch_{i_4}(N) \subsetneqq \Psi_{i_4}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right).$$

[3] Finally, if there is no agent satisfying either [1], nor [2], select i_4 as follows

[α] $i_1 \in Ch_{i_4}(N)$ if $i_3 \notin Ch_{i_1}(N)$, and there is no i_4 such that $i_2 \in Ch_{i_4}(N)$. Note that, in this case $i_3 \in Ch_{i_2}(N)$, $i_2 \in Ch_{i_1}(N)$, as explained in Condition (a_2) above, and $i_1 \in Ch_{i_4}(N)$. Henceforth, we have that

$$Ch_{i_4}(N) \nsubseteq \Psi_{i_4}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right)$$

because

$$\Psi_{i_3}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right) = \{i_3\},$$

$$\{i_2\} P^{tc}_{i_2} S \text{ for any } S \in \mathcal{A}^{i_2} \text{ not containing } i_3,$$

and

$$\{i_1\} P^{tc}_{i_1} S \text{ for any } S \in \mathcal{A}^{i_1} \text{ not containing } i_2,$$

or,

[β] $i_2 \in Ch_{i_4}(N)$ if $i_3 \notin Ch_{i_2}(N)$, and there is no i_4 such that $i_1 \in Ch_{i_4}(N)$. Note that, in this case $i_3 \in Ch_{i_1}(N)$, $i_1 \in Ch_{i_2}(N)$, and $i_2 \in Ch_{i_4}(N)$. Henceforth, we have that

$$Ch_{i_4}(N) \nsubseteq \Psi_{i_4}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right).$$

This is because

$$\Psi_{i_3}\left(R_{-\{i_1,i_2,i_3\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}\right) = \{i_3\},$$

$$\{i_1\} P^{tc}_{i_1} S \text{ for any } S \in \mathcal{A}^{i_1} \text{ not containing } i_3,$$

and

$$\{i_2\} P^{tc}_{i_2} S \text{ for any } S \in \mathcal{A}^{i_2} \text{ not containing } i_1,$$

Note that, since $i_4 \in \mu^{tc}_i(R) \setminus \{i_1, i_2, i_3\}$, we must be able to find i_4 to satisfy any of the three cases above.

Now, note that stability of $\Psi_{i_4}\left(R_{-\{i_1,i_2,i_3,i_4\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}, R^{tc}_{i_4}\right)$, will imply that

$$\Psi_{i_4}\left(R_{-\{i_1,i_2,i_3,i_4\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}, R^{tc}_{i_4}\right) = \{i_4\}, \text{ or}$$

$$Ch_{i_4}(N) \subseteq \Psi_{i_4}\left(R_{-\{i_1,i_2,i_3,i_4\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}, R^{tc}_{i_4}\right).$$

Henceforth, strategy-proofness for Ψ implies that $\{i_4\}$ must coincide with the agreements configuration $\Psi_{i_4}\left(R_{-\{i_1,i_2,i_3,i_4\}}, R^{tc}_{i_1}, R^{tc}_{i_2}, R^{tc}_{i_3}, R^{tc}_{i_4}\right)$.

We should consider, again the following two cases:

(a_4) $\mu_i^{tc}(R) = \{i_1, i_2, i_3, i_4\}$. Then,

$$\cup_{k=1}^{4} Ch_{i_k}(N) = \{i_1, i_2, i_3, i_4\},$$

which contradicts that $\Psi\left(R_{-\{i_1,i_2,i_3\}}, R_{i_1}^{tc}, R_{i_2}^{tc}, R_{i_3}^{tc}\right)$ was stable for the problem $\left\{N, \left(R_{-\{i_1,i_2,i_3\}}, R_{i_1}^{tc}, R_{i_2}^{tc}, R_{i_3}^{tc}\right)\right\}$, and

(b_4) $\mu_i^{tc}(R) \neq \{i_1, i_2, i_3, i_4\}$. Then, there should be an agent, i_5 in $\mu_i^{tc}(R) \setminus \{i_1, i_2, i_3, i_4\}$ such that $\{i_1, i_2, i_3, i_4\} \cap Ch_{i_5}(N) \neq \emptyset$.

We can now select agent i_5 following a priorities rule similar to the previous one (for agent i_4), i.e., first, select i_5 such that $i_4 \in Ch_{i_5}(N)$; or, if such an agent does not exist, and since $i_5 \in \mu_i^{tc}(R)$, it must be that $i_5 \in \cup_{k=1}^{3} Ch_{i_k}(N)$. Hence, we can replicate the arguments on the steps [2] and [3] above exhausting all the possibilities, and showing that $\Psi_{i_5}\left(R_{-\{i_1,i_2,i_3,i_4\}}, R_{i_1}^{tc}, R_{i_2}^{tc}, R_{i_3}^{tc}, R_{i_4}^{tc}\right)$ must be $\{i_5\}$.

Since $\mu_i^{tc}(R)$ has a finite number of agents, an inductive argument help us to recover this set:

$$\mu_i^{tc}(R) = \{i_1, \ldots, i_{r-1}, i_r\},$$

in such a way that agent i_r:

(a) Could manipulate Ψ at $\left(R_{-\{i_1,\ldots,i_{r-1}\}}, R_{i_1}^{tc}, \ldots, R_{i_{r-1}}^{tc}\right)$ via $R_{i_r}^{tc}$ or,

(b) $\Psi\left(R_{-\{i_1,\ldots,i_r\}}, R_{i_1}^{tc}, \ldots, R_{i_r}^{tc}\right)$ is not stable for $\left\{N, \left(R_{-\{i_1,\ldots,i_r\}}, R_{i_1}^{tc}, \ldots, R_{i_r}^{tc}\right)\right\}$,

which contradicts the initial hypothesis on Ψ's properties. ■

5. Conclusions

In this paper we have explored the existence of economic environments where the problem of reaching stable coalition structures can be solved satisfactorily. The property that characterizes this environment can be understood as a natural generalization of what Alcalde and Romero-Medina (2000) called Essentiality.

This generalization introduces a new feature to the core of the problems that we study. We show that the core is not only non-empty, but also that it is not always a singleton. This is why we explore the possibilities of agents' strategic behavior when they face mechanisms to select stable agreement configurations. We show the existence of mechanisms for which agents have no incentive to misinform about their true characteristics (or preferences), when they are restricted to satisfy the Tops Responsiveness condition. We also deal with the problem of characterizing the (non-empty) family of stable mechanisms. Our main result is therefore quite strong: Whenever a strategy-proof mechanism for selecting stable outcomes is required, we have a unique option, that of employing the *Tops Covering Mechanism*.

The above result contrasts with the findings of Sönmez (1999), who shows that in a more general setting, the existence of strategy-proof stable rules is conditioned to the case of frameworks whose core is (essentially) a singleton. Provided that, as shown by Example 4.4, when agents' preferences satisfy the Tops Responsiveness condition the core is not necessarily single-valued (in Sönmez's terms), we should provide arguments for such an apparent *contradiction*. Sönmez's result rests on a condition which might be considered strong in some environments, namely each agent can truncate her preferences at the level she wants to. In fact, Sönmez's conditions state, in particular, that given two alternatives, x and y, and an agent's preferences P_i if $xP_iyP_ia_i$ holds, there is another agent's preferences, say P_i' satisfying that $xP_i'a_iP_i'y$, where declaring $a_iP_i'y$ is to say that alternative y is considered individually irrational by agent i. Nevertheless, this condition is never satisfied by frameworks that satisfy the Tops Responsiveness condition. This is the reason why there is no contradiction between our result and the Sönmez's one, in contrast to what could seem apparently.

APPENDIX: *Proof of Theorem 3.4*

We shall now deal with the study of the Tops Covering Algorithm. We first show, Lemma 5.1, that it converges in a finite number of steps; we then show, Lemma 5.2, that our rule provides an agreements configuration for any problem; finally, we proceed to see that this agreement configuration is stable, relative to the agents' preferences.

Lemma 5.1. *Let $\{N, R\}$ be an agreements problem with agents' preferences that satisfy the Tops Responsiveness Condition. The Tops Covering Algorithm then converges in a finite number of steps.*

Proof. We first concentrate on showing that N^1 in Step (1) is non-empty. The proof is completed by an iterative argument on the following steps.

By following Definition 3.1, we can construct, for each agent i in N, the set $\tau_i(N)$. Let us assume that N^1 in Step (1) is empty. This means that, for each $i \in N$, there is some agent $i' \in \tau_i(N)$, such that $\tau_{i'}(N) \subset \tau_i(N)$. Let us construct the sequence of individuals $\{i_k\}_{k=0}^{\infty}$ such that $i_0 = 1$, and, for each $k \geq 1$, i_k satisfies that $\tau_{i_k}(N) \subset \tau_{i_{k-1}}(N)$. Since the set of agents is finite, this sequence must have a cycle. Henceforth, there must be two terms, say k and k', such that $i_k = i_{k'}$. By transitivity of the (strict) inclusion relation, we have that $\tau_{i_k}(N) \subset \tau_{i_k}(N)$. A contradiction. Henceforth, there should be an agent i such that $\tau_i(N) \subseteq \tau_h(N)$ for all $h \in \tau_i(N)$, i.e. N^1 is non-empty.

Now, let us consider the problem[7] $\left\{N \backslash N^1, R^{N \backslash N^1}\right\}$, where N^1 is the set described in the Step (1) of Definition 3.2. Note that, by using the above argument, we have that we could also find a non-empty set of agents in this problem, say M, playing the same role as N^1 does in $\{N, R\}$. Since $N \backslash N^1$ is the set of agents considered in Step (2) of Definition 3.2, we conclude that M is precisely the set N^2 in the description of the Tops Covering Algorithm. Since N is finite, an inductive argument on the set of agents yields the desired result. ■

[7] We follow the next notation. Given a problem $\{N, R\}$, and a set of agents S, let $\left\{S, R^S\right\}$ denote the problem whose agents are in S, and their preferences are restricted to S in a natural way; i.e., for any i in S and two sets of agents, say T and T' contained in S with $i \in T \cap T'$, we have that $T R_i^S T'$ if, and only if, $T R_i T'$.

Lemma 5.2. *Let $\{N, R\}$ be an agreements problem whose agents' preferences satisfy the Tops Responsiveness Condition. Then, the output of the Tops Covering Algorithm is an agreements configuration for $\{N, R\}$.*

Proof. Let us assume that the algorithm stops after t steps. Let us consider the sets N^k, $k = 1, \ldots, t$, that are produced at each step of the Tops Covering Algorithm. Let suppose that agent i belongs to N^1, we show that, in such a case $\mu_i^{tc}(R) \subseteq N^1$. Note that, by construction, it should be the case that, for all i' in $\mu_i^{tc}(R)$, $\tau_i(N) = \tau_{i'}(N)$. Since $\tau_i(N) \subseteq \tau_h(N)$ for each $h \in \tau_i(N)$, this property must also hold for agent i'. Hence, for all i' in $\tau_i(N)$, $\mu_{i'}^{tc}(R) = \mu_i^{tc}(R) = \tau_i(N)$. Considering that the second step of our algorithm can be described as the first step applied to the problem $\left\{N \backslash N^1, R^{N \backslash N^1}\right\}$, we can conclude that, for each i in N^2, and any i' in $\tau_i(N \backslash N^1)$, $\mu_{i'}^{tc}(R) = \mu_i^{tc}(R) = \tau_i(N \backslash N^1)$. Applying an inductive argument to the above statement, it is straightforward that, for any set N^k, $k > 1$, and each two agents i in N^k, and i' in $\tau_i\left(N \backslash \cup_{j=1}^{k-1} N^j\right)$ it should be the case that $\mu_{i'}^{tc}(R) = \mu_i^{tc}(R) = \tau_i\left(N \backslash \cup_{j=1}^{k-1} N^j\right)$. ■

To conclude this section, we shall show that, for any R satisfying the Tops Responsiveness Condition, the agreements configuration $\mu^{tc}(R)$ is stable for $\{N, R\}$.

Proposition 5.3. *Let $\{N, R\}$ be an agreements problem. If agents' preferences satisfy the Tops Responsiveness Condition, the Tops Covering Algorithm produces an agreement which is stable for $\{N, R\}$.*

Proof. Let the problem $\{N, R\}$ satisfy the Tops Responsiveness Condition, and let $\mu^{tc}(R)$ be the output of applying the Tops Covering Algorithm to $\{N, R\}$. Assume that $\mu^{tc}(R)$ is unstable, then there should be a non-empty set of agents, say S, such that, for each i in S, $SP_i\mu_i^{tc}(R)$.

For each agent i, let $k(i)$ denote the stage of the algorithm in which this agent reach her agreement[8] and let us define the set $N_i = N \backslash \cup_{1 \leq k < k(i)} N^k$. By construction, it should clearly be the case that, for each i in S, $Ch_i(N_i) \subseteq \mu_i^{tc}(R) \subseteq N_i$.

Let us select an agent $\hat{\imath}$ in S with the lower $k(\hat{\imath})$, i.e., for each i in S, $k(\hat{\imath}) \leq k(i)$. By Definition 2.2, and the fact that S blocks $\mu^{tc}(R)$, we have that $Ch_{\hat{\imath}}(N_{\hat{\imath}}) \subseteq S$, and $\mu_{\hat{\imath}}^{tc}(R) \nsubseteq S$. We must therefore consider two possibilities: First, that if S is a subset of $\mu_{\hat{\imath}}^{tc}(R)$, and hence $k(i) = k(\hat{\imath})$ for any agent i in S. Then, $Ch_S(N_{\hat{\imath}}) = S$, which contradicts the fact that $\tau_{\hat{\imath}}(N_{\hat{\imath}})$ were the smaller set containing $Ch_{\hat{\imath}}(N_{\hat{\imath}})$ such that $Ch_{\tau_{\hat{\imath}}(N_{\hat{\imath}})}(N_{\hat{\imath}}) = \tau_{\hat{\imath}}(N_{\hat{\imath}})$. The second possibility

[8] This formulation was introduced in the proof of Theorem 4.3.

is that there exists an agent, say i' in $S \cap \mu_{\hat{\imath}}^{tc}(R)$ such that $Ch_{i'}(N_{\hat{\imath}}) \nsubseteq S$. This implies that $\mu_{i'}^{tc}(R)\, P_{i'} S$, which contradicts the conclusion that i' belonged to the blocking coalition S. ■

Note that the fact stated in Theorem 3.4 is a direct consequence of the above proposition.

REFERENCES

ALCALDE, J. (1995). "Exchange-Proofness or Divorce-Proofness? Stability in One-Sided Matching Markets." *Economic Design* **1**, 275-287.

ALCALDE, J., BARBERÀ, S. (1994). "Top Dominance and the Possibility of Stable Rules for Matching Problems," *Economic Theory* **4**, 417-425.

ALCALDE, J., ROMERO-MEDINA, A. (2000). "Coalition Formation and Stability," WP 00–31, Universidad Carlos III de Madrid.

BANERJEE, S., KONISHI, H. and SÖNMEZ, T. (2001). "Core in a Simple Coalition Formation Game," *Social Choice and Welfare* **18**, 135-153.

BOGOMOLNAIA, A., JACKSON, M. (1998). "The Stability of Hedonic Coalition Structures," *Games and Economic Behavior*, forthcoming.

CECHLÁROVÁ, K., ROMERO-MEDINA, A. (1998). "Stability in One-Sided Matching Markets," working paper 98-52. Universidad Carlos III de Madrid.

DRÈZE, J., GREENBERG, J. (1980). "Hedonic Coalitions: Optimality and Stability," *Econometrica* **48**, 987-1003.

GALE, D., SHAPLEY, L.S. (1962). "College Admissions and the Stability of Marriage," *American Mathematical Monthly* **69**, 9-15.

GIBBARD, A. (1973). "Manipulation of Voting Schemes: A General Result," *Econometrica* **41**, 587-601.

ROTH, A.E. (1982). "The Economics of Matching: Stability and Incentives," *Mathematics of Operations Research* **7**, 617-628.

ROTH, A.E., SOTOMAYOR, M. (1990). Two-sided Matching: A Study in Game-Theoretic Modeling and Analysis. New York: Cambridge University Press.

SATTHERWAITE, M. (1975). "Strategy-Proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions," *Journal of Economic Theory* **10**, 187-216.

SHAPLEY, L.S., SCARF, H. (1974). "On Cores and Indivisibility," *Journal of Mathematical Economics* **1**, 23-28.

SÖNMEZ, T. (1999). "Strategy-Proofness and Essentially Single-Valued Cores," *Econometrica* **67**, 677-690.