Coalition Formation with Size Constraints and Quality Effect

Claudia Meo* Antonio Miralles Asensio† Pietro Salmaso‡ Duygu Sili§

Abstract

We study a coalition formation model in a purely hedonic setting in which the number of coalitions and their sizes are fixed *ex ante*. We establish the existence of swap–stable partitions when the utility w_{ij} that agent i derives from being in the same coalition as agent j is the sum of two components: a symmetric "friendship" term v_{ij}, in the spirit of Bogomolnaia and Jackson (2002), and an agent-specific coefficient t_j capturing the "quality effect"—either beneficial or disruptive—that agent j exerts on the coalition as a whole. Our result generalizes that of Bilò et al. (2022) to non–symmetric preferences.

1 Introduction

We study a model of coalition formation in purely hedonic settings[1] where agents' preferences depend solely on the composition of their own coalition, and external effects across coalitions are absent.

*Department of Economics and Statistics, University of Naples Federico II, Via Cintia, 80126 Naples, Italy. claudia.meo@unina.it

†Department of Economics, University of Messina, Piazza Pugliatti, 98122 Messina, Italy. antonio.mirallesasensio@unime.it

‡Department of Economics and Statistics, University of Naples Federico II, Via Cintia, 80126 Naples, Italy. pietro.salmaso@unina.it

§Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, 28903 Getafe (Madrid), Spain. dsili@eco.uc3m.es

[1] The term *hedonic* was first introduced by Drèze and Greenberg (1980), who used it to describe coalition formation settings in which each agent's preferences depend solely on the composition of the group they belong to.

Unlike standard formulations, we impose that both the number of coalitions and their sizes are fixed *ex ante*. This assumption is motivated by a variety of real-world applications, such as the assignment of students to classes or researchers to project teams. Fixing coalition sizes renders several classical stability notions inapplicable, as they rely on deviations that alter coalition composition. In particular, core stability (see, Banerjee et al. 2001, who provide sufficient conditions for the existence of core stable partitions such as the top coalition and weak top coalition properties), as well as Nash and individual stability (see, Bogomolnaia and Jackson 2002, who establish existence under additively separable and symmetric preferences), are no longer suitable in this setting.

To address this limitation, we adopt the notion of swap stability, introduced by Bilò et al. (2022). A partition is swap-stable if no pair of agents can profitably exchange coalitions. Bilò et al. (2022) establish the existence of swap–stable partitions in additively separable hedonic games with symmetric preferences.

We extend their framework by relaxing the symmetry of preferences. Specifically, we assume that the utility w_{ij} that agent i derives from being in the same coalition as agent j is the sum of two terms: a symmetric "friendship" value v_{ij}, as in Bogomolnaia and Jackson (2002), and an agent-specific coefficient t_j capturing the "quality effect"—either beneficial or disruptive—that agent j exerts on the coalition as a whole. Although the introduction of this term generally induces asymmetries in preferences, we show that a swap–stable partition still exists under our utility specification. This result highlights the robustness of swap stability to richer and more realistic preference structures and stands in contrast to the nonexistence of swap-stable partitions when friendship values are not symmetric, as shown in Bilò et al. (2022).

The proposed model is particularly well suited to class formation problems, in which students are assigned to classes of predetermined size and both their preferences over classmates and their heterogeneous impact on the learning environment are taken into account. Our framework naturally incorporates both aspects and provides a solid theoretical basis for analyzing stability in such settings.

2 The model

We consider a finite set N of n agents, a number of coalitions $k \in \mathbb{N}$, and a capacity vector $q = (q_1, \ldots, q_k) \in \mathbb{N}^k$. A k-partition of N with size constraints is a partition $P = \{C_1, \ldots, C_k\}$ of N into k coalitions such that $|C_i| = q_i$ for each $i \in \{1, \ldots, k\}$. We assume that $\sum_{i=1}^k q_i = n$.

Each agent i has a linear order $\succeq_i$ over the set $\{S_t : S_t \subset N \text{ and } i \in S_t\}$. We assume that $\succeq_i$ is additively separable; that is, there exists a function $w_i : N \to \mathbb{R}$ such that for all $S_1, S_2 \subset N$ with $i \in S_1 \cap S_2$:

$$S_1 \succeq_i S_2 \Leftrightarrow \sum_{j \in S_1} w_i(j) \geq \sum_{j \in S_2} w_i(j)$$

We also assume that, for each $i \in N$, $w_i = v_i + t$ where $v_i : N \to \mathbb{R}$ is the friendship component while $t : N \to \mathbb{R}$ represents the quality effect; $t(j)$ captures the positive or negative impact of agent $j \in N$ on the other agents. These values may stem from an exogenous and commonly accepted ranking of agents, reflecting an objective assessment of their individual quality. Without loss of generality, we assume that, for every $i \in N$, $v_i(i) = 0$ and v_i is symmetric, that is, $v_i(j) = v_j(i)$ for all $j \in N$. For ease of notation, we denote $w_i(j)$, $v_i(j)$ and $t(j)$ by w_{ij}, v_{ij} and t_j, respectively. By symmetry, $v_{ij} = v_{ji}$; note that, in general, $w_{ij} \neq w_{ji}$.

Given an agent i and a coalition C such that $i \in C$, the utility that agent i obtains from belonging to C is

$$U_i(C) = \sum_{j \in C} w_{ij}.$$

For a k–partition $P = \{C_1, \ldots, C_k\}$ of N, let

$$U_i(P) = U_i(C_h),$$

where C_h is the unique coalition in P such that $i \in C_h$.

A hedonic game G with size constraint and quality effect is a 5−tuple $G = (N, k, q, (v_i)_i, t)$ where N is the set of agents, k is the number of coalitions in the partition, q is a capacity vector, v_i is the friendship function of agent i and t is a quality function.

We focus on the notion of swap stability introduced by Bilò et al. (2022).

A k-partition $P = \{C_1, \ldots, C_k\}$ of N is *swap-stable* if there does not exist a pair of agents who would both strictly benefit from exchanging their current coalitions.

More generally, given a subset $S \subseteq N$, a partition P is *swap-stable for S* if there is no pair of agents in S who would both strictly benefit from swapping their assigned coalitions.

Bilò et al. (2022) show that if the functions v_i are symmetric, a swap-stable partition always exists and can be obtained by starting from an arbitrary partition and iteratively allowing improving swaps. They also provide an example demonstrating that the set of swap-stable partitions may be empty when the symmetry assumption on the coefficients w_{ij} is dropped.

3 Main result

We weaken the symmetry assumption of the paper of Bilò et al. (2022) and prove a stronger existence result of swap–stable partitions for a game with size constraints and quality effect.

Theorem 1. *For every hedonic game with size constraints and quality effect $G = (N, k, q, (v_i)_i, t)$, there exists a swap–stable partition.*

Proof. For any coalition C and any partition $P = \{C_1, \ldots, C_k\}$, define their potentials as $\Phi(C) = \sum_{i \in C} U_i(C)$ and $\Phi(P) = \sum_{h=1}^{k} \Phi(C_h)$, respectively. It holds that:

$$\Phi(C) = \sum_{i \in C} U_i(C) = \sum_{(i,j) \in C \times C} w_{ij} = \sum_{(i,j) \in C \times C} (v_{ij} + t_j) = \sum_{i \in C} \sum_{j \in C} (v_{ij} + t_j).$$

Since the functions v_i are symmetric, this expression can be rearranged as

$$\Phi(C) = 2 \sum_{\{i,j\} \subseteq C} v_{ij} + |C| \sum_{i \in C} t_i.$$

Claim. Suppose that q is a constant vector, that is $q_h = c$ for all $h = 1, \ldots, k$. Let $P = \{C_1, \ldots, C_k\}$

be a partition with $|C_i| = c$ for every $i = 1, \ldots, k$ and let $S \subseteq N$. Then there exists a partition $P' = \{C'_1, \ldots, C'_k\}$ that is swap-stable for S and such that, for every agent $i \in N \setminus S$ and every $h \in \{1, \ldots, k\}$,

$$i \in C'_h \quad \text{if and only if} \quad i \in C_h.$$

Proof. We first show that every swap strictly increases the potential of the partition.

Let $P = \{C_1, \ldots, C_k\}$ be a partition and suppose that two agents $i, j \in N$ swap coalitions. Without loss of generality, assume that $i \in C_1$ and $j \in C_2$.

Define

$$\overline{C}_1 = C_1 \setminus \{i\}, \qquad \overline{C}_2 = C_2 \setminus \{j\},$$

and let

$$P' = \{C'_1, \ldots, C'_k\} = \{\{j\} \cup \overline{C}_1,\ \{i\} \cup \overline{C}_2,\ C_3, \ldots, C_k\}.$$

Since the swap is improving,

$$U_i(P') > U_i(P) \quad \text{and} \quad U_j(P') > U_j(P).$$

We now prove that

$$\Phi(P') - \Phi(P) > 0.$$

Only coalitions C_1 and C_2 are affected by the swap. Hence,

$$\Phi(P') - \Phi(P) = \Phi(C'_1) - \Phi(C_1) + \Phi(C'_2) - \Phi(C_2).$$

Since $|C_1| = |C_2| = c$, we obtain after simplification:[2]

[2]The terms $|C|\ \sum_{i \in C} t_i$ cancel out only because we assumed in this claim that $|C_h| = c$ for every $h = 1, .., k$

$$\Phi(P') - \Phi(P) = 2\sum_{m\in\overline{C}_1} v_{jm} - 2\sum_{m\in\overline{C}_1} v_{im} + 2\sum_{m\in\overline{C}_2} v_{im} - 2\sum_{m\in\overline{C}_2} v_{jm}.$$

Rearranging terms yields

$$\Phi(P') - \Phi(P) = 2\Bigg[\Bigg(\sum_{m\in\overline{C}_2} v_{im} - \sum_{m\in\overline{C}_1} v_{im}\Bigg) + \Bigg(\sum_{m\in\overline{C}_1} v_{jm} - \sum_{m\in\overline{C}_2} v_{jm}\Bigg)\Bigg].$$

The first bracket coincides with the gain in utility of agent i, while the second bracket coincides with the gain in utility of agent j. Since the swap strictly improves utilities of both agents, we get:

$$\Phi(P') - \Phi(P) = 2\big[U_i(P') - U_i(P) + U_j(P') - U_j(P)\big] > 0.$$

Now, let $P = \{C_1, \ldots, C_k\}$ be any initial partition and let $S \subseteq N$. Construct a sequence of partitions $(P_t)_{t\geq 0}$ as follows. Set $P_0 = P$. If, at step t, there exists a pair of agents in S who both strictly benefit from swapping coalitions under P_t, define P_{t+1} as the partition obtained from P_t after such a swap. If no such pair exists, the process stops.

Since the potential strictly increases after each improving swap and the number of partitions is finite, the process cannot cycle and must terminate in finitely many steps. The terminal partition is swap-stable for S. Moreover, only agents in S are allowed to swap, so the assignment of agents in $N \setminus S$ remains unchanged throughout the process.

Consider now the case of heterogeneous capacities. Let $q = (q_1, \ldots, q_k)$ be the capacity vector and let

$$Q = \max_h q_h.$$

Construct an auxiliary problem where coalitions have the same capacities as follows. Let

$$N' = N \cup D,$$

where D is a set of dummy agents and $|D| = kQ - n$. Dummy agents generate no utility and no externalities; that is, for every $d \in D$ and every $i \in N'$, $t_d = 0$ and $v_{id} = 0$.

Consider any partition $P = \{C_1, \dots, C_k\}$ of N' such that, for each h, coalition C_h contains exactly $Q - q_h$ dummy agents. By the previous claim, there exists a partition $P' = \{C'_1, \dots, C'_k\}$ of N' that is swap-stable for N and such that, for every h, coalition C'_h still contains exactly $Q - q_h$ dummy agents.

Define

$$P'' = \{C'_1 \cap N, \dots, C'_k \cap N\}.$$

We claim that P'' is swap-stable for the original problem.

Observe that for every h and every $i \in C'_h \cap N$,

$$U_i(C'_h) = U_i(C'_h \cap N),$$

since dummy agents contribute zero to utility of the other agents.[3]

If a pair of agents in N could profitably swap in P'', the same pair would profitably swap in P' in the auxiliary problem, contradicting the swap-stability of P' for N. Therefore, P'' is swap-stable in the original problem. □

It is easily checked that the theorem continues to hold when, in the definition of swap stability, we require that one of the two agents strictly improves while the other does not become worse off.

The following example shows that swap–stable partitions might not exist when there is overcapacity, that is $\sum_{i=1}^{k} q_i \geq n$ and agents can not only swap, but occupy an empty place in a class that is not completely filled.

Example 1. Let $N = \{1, 2, 3\}$, $k = 2$ and $q_1 = q_2 = 2$. Assume that $t_1 = t_3 = v_{13} = 1$, $t_2 = 5$, $v_{23} = -2$, $v_{12} = -4$. There are three possible 2–partitions of N: $P_1 \equiv \{\{1, 2\}, \{3\}\}$,

[3]However, dummy agents are utility receivers and this explains why in the original problem it is not true that the potential of the partition increases after every swap.

$P_2 \equiv \{\{1,3\},\{2\}\}$, $P_3 \equiv \{\{2,3\},\{1\}\}$. Let us show that none of them is swap–stable.

a. $U_1(P_1) = 1 < 2 = U_1(P_2)$ so agent 1 would leave the coalition with agent 2 and occupy the empty place in coalition with agent 3. P_1 is not swap stable.

b. $U_3(P_3) = 3 > 2 = U_3(P_2)$ so agent 3 would leave the coalition with agent 1 and occupy the empty place in coalition with agent 2. P_2 is not swap stable.

c. $U_1(P_2) = 2 > 0 = U_1(P_3)$ and $U_2(P_3) = -1 < 0 = U_2(P_2)$, so agents 1 and 2 would swap their places in P_3. P_3 is not swap stable.

References

Banerjee, S., H. Konishi, and T. Sönmez 2001. Core in a Simple Coalition Formation Game. Social Choice and Welfare, 18, 135–153.

Bilò, V., G. Monaco and L. Moscardelli 2022. Hedonic games with fixed size coalitions. AAAI-22, 9287–9295.

Bogomolnaia, A. and M. Jackson 2002. The Stability of Hedonic Coalition Structures. Games and Economic Behavior, 38, 201–230.

Drèze, J. and J. Greenberg 1980. Hedonic Coalitions: Optimality and Stability. Econometrica, 48, 987–1003.