yönetim ve ekonomi araştırmaları dergisi
journal of management and economics research
Cilt/Volume: 20 Sayı/Issue: 4 Aralık/December 2022 ss. /pp. 335-350
M. Karakaya, S. Özbilen http://dx.doi.org/10.11611/yead.1138933

INTERNAL STABILITY AND PARETO OPTIMALITY IN HEDONIC COALITION FORMATION GAMES

Asst. Prof. (Ph.D.) Mehmet KARAKAYA*

Asst. Prof. (Ph.D.) Seçkin ÖZBİLEN**

ABSTRACT

We study hedonic coalition formation games that consist of a finite set of agents and a list of agents' preferences such that each agent's preferences depend only on the members of her coalition. An outcome of a hedonic coalition formation game is a partition (i.e., coalition structure) of the finite set of agents. We study the existence of partitions that are both internally stable and Pareto optimal. We construct an algorithm that terminates for each given hedonic coalition formation game such that the outcome of the algorithm is internally stable and Pareto optimal. We also show that if the outcome of the algorithm is the partition that consists of singleton coalitions then it is also core stable and if it is the partition that contains only the grand coalition then it is also both core stable and Nash stable.

Keywords: *Coalition Formation, Core Stability, Hedonic Games, Internal Stability, Nash Stability, Pareto Optimality.*

JEL Classification: *C71, C78, D71.*

HEDONİK KOALİSYON OLUŞUM OYUNLARINDA İÇSEL KARARLILIK VE PARETO VERİMLİLİK

ÖZET

Bu makalede hedonik koalisyon oluşum oyunları çalışılmıştır. Bir hedonik koalisyon oluşum oyunu sonlu sayıda bireyler kümesi ile bu bireylerin tercih listesinden oluşur öyle ki her bireyin tercihi kendisinin üyesi olabileceği bütün koalisyonların bir sıralamasıdır. Bir hedonik koalisyon oluşum oyununun çıktısı sonlu bireyler kümesinin bir partisyonudur (koalisyon yapısıdır). Bu makalede içsel kararlı ve Pareto verimli partisyonların varlığını araştırıyoruz. Bir algoritma tanımlıyoruz öyle ki bu algoritma verilen her hedonik koalisyon oluşum oyunu için duruyor ve durduğu zaman içsel kararlı ve Pareto optimal bir partisyonu çıktı olarak veriyor. Ayrıca, bu algoritmanın çıktısı tekil koalisyonlardan oluşan partisyon ise bunun aynı zamanda çekirdek kararlı olduğunu, eğer çıktı sadece büyük koalisyonu

* İzmir Katip Çelebi University, Department of Economics, İzmir, Türkiye, e-mail: mehmet.karakaya@ikcu.edu.tr.
** Özyeğin University, Faculty of Business, İstanbul, Türkiye, e-mail: seckin.ozbilen@ozyegin.edu.tr.

Makale Geçmişi/Article History

Başvuru Tarihi / Date of Application	: 12 Ağustos/ August 2022
Düzeltme Tarihi / Revision Date	: 20 Ekim / October 2022
Kabul Tarihi / Acceptance Date	: 3 Aralık / December 2022

Araştırma Makalesi/Research Article

içeren partisyon ise bunun aynı zamanda hem çekirdek kararlı hem de Nash kararlı olduğunu gösteriyoruz.

***Anahtar Kelimeler**: Koalisyon Oluşumu, Çekirdek Kararlılık, Hedonik Oyun, İçsel Kararlılık, Nash Kararlılık, Pareto Verimlilik.*

***JEL Kodları**: C71, C78, D71.*

1. INTRODUCTION

We see many instances of coalition formation in our daily lives: college students form study or project groups among themselves in a course, faculty members in a department form research groups, people form charities and clubs, and politicians form new parties or alliances to win elections. Most of these coalition formation instances can be modeled as a hedonic coalition formation game.

A hedonic coalition formation game (in brief, a hedonic game) consists of a finite set of agents and a list of agents' preferences such that each agent has preferences over all coalitions each of which contains her. That is, each agent's preferences depend only on the identity of members of her coalition. This hedonic aspect of preferences was firstly introduced by Drèze and Greenberg (1980) and the formal model of hedonic coalition formation games was introduced by Banerjee, Konishi and Sönmez (2001) and Bogomolnaia and Jackson (2002). Hedonic coalition formation games are also called as simple coalition formation problems or coalition formation problems without externalities. Examples are the formation of clubs and organizations, provision of public goods in local communities, and forming research groups. Moreover, marriage problems and roommate problems (Gale and Shapley, 1962; Roth and Sotomayor, 1990) are special hedonic coalition formation games at which a coalition consists of at most two agents.

An outcome of a hedonic game is a partition, i.e., a collection of pairwise disjoint coalitions such that their union is equal to the set of agents. Some stability and efficiency properties of partitions are studied in the literature.[1] Mostly studied stability concepts are Nash stability and core stability, and the efficiency concept is Pareto optimality.

Given a partition for a hedonic game, an agent Nash blocks this partition if it is beneficial for her to move another coalition of the partition, i.e., she leaves her current coalition and joins another coalition of the partition or to the empty set and by this movement she is better off. A partition for a hedonic game is Nash stable if it is not Nash blocked by any agent.[2] Note that Nash stability is based on an individual deviation among the coalitions of a given partition.

[1] See Sung and Dimitrov (2007) for the taxonomy of some stability concepts.
[2] Weaker versions of Nash stability are individual stability, contractual Nash stability, and contractual individual stability. For more information about these stability notions, we refer interested reader to Bogomolnaia and Jackson (2002), Sung and Dimitrov (2007), Pápai (2007), and Suksompong (2015).

A partition for a hedonic game is core stable if there is no coalition such that each agent in this coalition prefers it to her current coalition under the partition. If such a coalition exists we then say that this coalition blocks the partition.[3] Note that core stability is based on a coalitional deviation such that each agent in the blocking coalition leaves her current coalition and then they form a new coalition. So, the coordination of agents who are in different coalitions are needed for a blocking coalition. However, such a coordination can be impossible for some cases. Then, for those cases internal stability is an appropriate stability notion.[4]

A coalition is internally stable if no group of agents within the coalition be better off by leaving the coalition and forming a new coalition on their own. A partition for a hedonic game is internally stable if each coalition in the partition is internally stable. Note that internal stability requires the coordination of agents who are in the same coalition, hence internal stability is an appropriate stability notion when coordination among coalitions is not possible. So, internal stability is a weaker stability notion than core stability, that is, a core stable partition is internally stable but an internally stable partition may not be core stable. We also note that each hedonic game has an internally stable partition, the partition that consists of singleton coalitions is internally stable. The concepts of the internally stable coalition and internally stable partition were first introduced by Dimitrov, Borm, Hendrickx and Sung (2006) and Alcalde and Romero-Medina (2006). Dimitrov et al. (2006) proved the existence of core stable partitions in hedonic games satisfying the appreciation of friends or aversion of enemies properties by using the concepts of the internally stable coalition and internally stable partition[5]. Alcalde and Romero-Medina (2006) proved the existence of core stable partitions in the set of hedonic games that satisfy the intersection responsiveness property by using the concept of the internally stable coalition[6]. Liu, Tang and Fang (2014) adopted the concept of internal stability to the concepts of matching and exchange. Schlueter and Goldsmith (2020) investigated the relationship between the concepts of internal stability and Nash stability in different domains of hedonic games. Özbilen (2022) analyzed the relation between internal stability and other stability concepts and introduced an algorithm that brings all internally stable partitions in the domain of all hedonic games.

[3] The existence of core stable or Nash stable partitions is not guaranteed in general. The existence of such partitions is guaranteed under some conditions. For conditions guaranteeing the existence of a core stable partition, see Banerjee et al. (2001), Bogomolnaia and Jackson (2002), Cechlárová and Romero-Medina (2001), Burani and Zwicker (2003), Alcalde and Revilla (2004), Alcalde and Romero-Medina (2006), Dimitrov, Borm, Hendrickx and Sung (2006), and Suzuki and Sung (2010). For conditions guaranteeing the existence of a unique core stable partition, see Pápai (2004) and İnal (2019). For a necessary and sufficient condition guaranteeing the existence of a core stable partition, see Iehlé (2007). For conditions guaranteeing the existence of a Nash stable partition, see Bogomolnaia and Jackson (2002), Burani and Zwicker (2003), Dimitrov and Sung (2004), Dimitrov and Sung (2006), Pápai (2007), and Suksompong (2015).
[4] There are studies related to stability concepts that are stronger or weaker than core stability or Nash stability. We refer the reader to Hajduková (2006), Karakaya (2011), Aziz and Brandl (2012), Aziz and Savani (2016), and Özbilen (2019) for more detailed analysis and discussions.
[5] For conditions of appreciation of friends and aversion to enemies we refer to Dimitrov and Sung (2004) and Dimitrov et al. (2006).
[6] For the intersection responsiveness property, see Alcalde and Romero-Medina (2006).

A partition for a hedonic game is Pareto optimal if there is no other partition at which all agents are weakly better off and some agents are strictly better off, i.e., it is impossible to make some agents strictly better off without hurting other agents. Bogomolnaia and Jackson (2002) showed that a Pareto optimal partition always exists for each hedonic game, and Aziz, Brandt and Harrenstein (2013) studied the computation and verification of the existence of Pareto optimal partitions in several sub-domains of hedonic games. Internal stability and Pareto optimality are independent of each other, i.e., an internally stable partition may not be Pareto optimal and a Pareto optimal partition may not be internally stable.

In this paper, we focus on internal stability and Pareto optimality of a partition. We construct an algorithm that terminates for each hedonic game and the outcome of the algorithm is internally stable and Pareto optimal (Theorem 1). Hence, an internally stable and Pareto optimal partition always exists. We also analyze the outcome of the algorithm when it is the partition consisting of only singleton coalitions and the partition consisting of the grand coalition in which all agents are together. If the outcome is the partition that consists of only singleton coalitions then it is also core stable (Proposition 1), and if the outcome is the partition that consisting of the grand coalition then it is also core stable and Nash stable (Proposition 2).

The rest of the paper is organized as follows. In Section 2, we present the hedonic coalition formation model and provide formal definitions of stability concepts and of Pareto optimality. In Section 3, we introduce our algorithm and present our results. Section 4 is dedicated to the conclusion and further discussions.

2. HEDONIC COALITION FORMATION MODEL

In this section, we introduce the hedonic coalition formation model and the stability concepts (Nash stability, core stability, and internal stability), efficiency concept (Pareto optimality) and voluntary participation (individual rationality) concept that we use throughout the paper.

Let $N = \{1,2,\dots,n\}$ be a finite set of agents with $n \geq 2$. A nonempty subset H of N is called a *coalition of* N. For each agent $i \in N$, let $\Sigma_i^N = \{S \subseteq N | i \in S\}$ denote the set of all coalitions of N containing agent i.

Each agent $i \in N$ has complete, transitive, and antisymmetric preferences $\succeq_i$ over Σ_i^N,[7] i.e., each agent has strict preferences over all coalitions of N each of which contains her. For an agent $i \in N$ and coalitions $S, T \in \Sigma_i^N$, $S \succ_i T$ means that agent i *strictly prefers* coalition S to coalition T, and $S \succeq_i T$ means that $S \succ_i T$ or $S = T$ that agent i *(weakly) prefers* coalition S to coalition T. We note that since

[7] A preference relation of agent i, $\succeq_i$, over Σ_i^N satisfies *completeness* if for each $S, T \in \Sigma_i^N$, $S \succeq_i T$ or $T \succeq_i S$, it satisfies *transitivity* if for each $S, T, H \in \Sigma_i^N$, if $S \succeq_i T$ and $T \succeq_i H$, then $S \succeq_i H$, and it satisfies *antisymmetry* if for each $S, T \in \Sigma_i^N$, $S \succeq_i T$ and $T \succeq_i S$ imply $S = T$.

each agent $i \in N$ has strict preferences over Σ_i^N, for each $S, T \in \Sigma_i^N$ with $S \neq T$ we have either $S \succ_i T$ or $T \succ_i S$.

A coalition $S \subseteq N$ with $i \in S$ is *acceptable* for agent i if $S \succeq_i \{i\}$ and it is *unacceptable* for agent i if $\{i\} \succ_i S$. For an agent $i \in N$ and her preferences $\succeq_i$, we denote her first-ranked (best) coalition by $r_1(i, \succeq_i)$, the second-ranked (second-best) coalition by $r_2(i, \succeq_i)$, and in a similar way, the k^{th}-ranked coalition by $r_k(i, \succeq_i)$. For instance, for an agent $i \in N$ and $S, T, H \in \Sigma_i^N$, $[\succeq_i : S \succ_i T \succ_i H \succ_i \{i\} \succ_i \ldots]$[8] means that agent i would first like to be a member of coalition S since $r_1(i, \succeq_i) = S$, then agent i would like to be a member of coalition T (since $r_2(i, \succeq_i) = T$) and then of coalition H (since $r_3(i, \succeq_i) = H$), etc.

For each $i \in N$, let $\mathcal{R}_i(\Sigma_i^N)$ denote the set of all strict preferences of agent i over Σ_i^N, and $\mathcal{R}^N = \prod_{i \in N} \mathcal{R}_i(\Sigma_i^N)$ denote the set of all preference profiles of agents in N. A list of agents' preferences $\succeq = (\succeq_1, \ldots, \succeq_n) \in \mathcal{R}^N$ is called a preference profile.

The hedonic aspect of preferences was firstly introduced by Drèze and Greenberg (1980) and the formal model of hedonic coalition formation games was introduced by Banerjee, Konishi and Sönmez (2001) and Bogomolnaia and Jackson (2002).

Definition 1 ***Hedonic Coalition Formation Game***

A hedonic coalition formation game, or a hedonic game, consists of a finite set of agents $N = \{1, 2, \ldots, n\}$ and their preferences $\succeq = (\succeq_1, \ldots, \succeq_n) \in \mathcal{R}^N$ and is denoted by $G = (N, \succeq)$.

An outcome of a hedonic game is a partition (coalition structure), that is, it is a collection of pairwise disjoint coalitions such that their union is equal to the set of agents.

Definition 2 ***Partition***

A partition π for a hedonic game $G = (N, \succeq)$ is a set $\pi = \{H_1, H_2, \ldots, H_K\}$ ($K \leq |N|$ is a positive integer) such that (i) for each $k \in \{1, \ldots, K\}$, $H_k \neq \emptyset$, (ii) $\bigcup_{k=1}^{K} H_k = N$, and (iii) for each $k, l \in \{1, \ldots, K\}$ with $k \neq l$, $H_k \cap H_l = \emptyset$.

For a partition π and an agent $i \in N$, we let $\pi(i)$ denote the unique coalition in π that contains agent i. The set of all partitions for a hedonic game $G = (N, \succeq)$ is denoted by $\Pi(N, \succeq)$ (even though the set of partitions is not related to the preference profile).

Since each agent in a hedonic game only cares about her own coalition, we extend an agent i's preferences $\succeq_i$ over Σ_i^N to over all partitions in a usual way (by using the same notation for preferences

[8] We list only acceptable coalitions in an agent's preferences, and "..." means that remaining unacceptable coalitions are ordered in any way. Here, coalitions S, T, H, and $\{i\}$ are acceptable for agent i and any other coalition is unacceptable.

over coalitions and partitions): for each agent $i \in N$ and partitions π, π', we have $\pi \succeq_i \pi'$ if and only if $\pi(i) \succeq_i \pi'(i)$.

We now define some properties of a partition. We start with the classical voluntary participation concept, namely the definition of individual rationality.

Definition 3 ***Individual Rationality***

A partition π is individually rational (IR) for hedonic game $G = (N, \succeq)$ if for each $i \in N$, $\pi(i) \succeq_i \{i\}$, i.e., each agent i finds her coalition $\pi(i)$ acceptable. Let $IR(N, \succeq)$ denote the set of all individually rational partitions for the hedonic game $G = (N, \succeq)$.

We define a well-known efficiency condition Pareto optimality. A partition is Pareto optimal if there is no other partition at which no agent is worse off and some agents are strictly better off.

Definition 4 ***Pareto Optimality***

A partition π is Pareto optimal (PO) for hedonic game $G = (N, \succeq)$ if there does not exist another partition $\pi' \in \Pi(N, \succeq)$ such that for each $i \in N, \pi' \succeq_i \pi$ and for some $j \in N, \pi' \succ_j \pi$. If such a partition π' exists then we say that π' Pareto dominates π. Let $PO(N, \succeq)$ denote the set of all Pareto optimal partitions for the hedonic game $G = (N, \succeq)$.

We now define Nash stability of a partition. As we noted in the Introduction that under Nash blocking an agent is allowed to move among the coalitions of a given partition. The existence of a Nash stable partition for hedonic coalition formation games under different conditions was firstly analyzed by Bogomolnaia and Jackson (2002).

Definition 5 ***Nash Stability***

We say that a partition π is Nash stable (NS) for hedonic game $G = (N, \succeq)$ if for each $i \in N$ and each $H \in (\pi \cup \{\emptyset\}), \pi(i) \succ_i H \cup \{i\}$. If such an agent $i \in N$ exists then we say that agent i Nash blocks π. Let $NS(N, \succeq)$ denote the set of all Nash stable partitions for the hedonic game $G = (N, \succeq)$.

We now define core stability for hedonic games. The existence of a core stable partition for hedonic coalition formation games under different conditions was firstly analyzed by Banerjee, Konishi and Sönmez (2001).

Definition 6 ***Core Stability***

A partition π is core stable (CS) for hedonic game $G = (N, \succeq)$ if there does not exist a coalition $H \subseteq N$ such that for each $i \in H$, $H \succ_i \pi(i)$. If such a coalition H exists then we say that coalition H blocks π. Let $CS(N, \succeq)$ denote the set of all core stable partitions for the hedonic game $G = (N, \succeq)$.

We note that for a coalition H to block a partition π, it requires the coordination of agents in H in a way that each agent $i \in H$ leaves her current coalition and then they form a new, self-standing coalition

H. However, when there is impossibility or a restriction of communication and hence coordination among the members of different coalitions, then core stability may not be an appropriate stability notion. Hence, when the communication among coalitions of a partition is impossible, a weaker stability concept than the core stability is appropriate to consider.

We now define internal stability for hedonic games which is a weaker stability notion than core stability. The concepts of the internally stable coalition and internally stable partition were first introduced by Dimitrov, Borm, Hendrickx and Sung (2006) and Alcalde and Romero-Medina (2006).

We say that a coalition $H \subseteq N$ is *internally stable* for hedonic game $G = (N, \succeq)$ if there does not exist a subset T of H such that for all $i \in T$, $T \succ_i H$. If such a coalition T of H exists then we say that T *internally blocks* H.

Definition 7 ***Internal Stability***

A partition π is internally stable (IntS) for hedonic game $G = (N, \succeq)$ if each coalition $H \in \pi$ is internally stable. Let $IntS(N, \succeq)$ denote the set of all internally stable partitions for hedonic game $G = (N, \succeq)$.

The relations between stability concepts and of Pareto optimality are given below, where ”⇒” indicates that if a partition satisfies the first notion then it also satisfies the second.

- $CS \Rightarrow IntS \Rightarrow IR$, core stability implies internal stability, and internal stability implies individual rationality.

- $CS \nRightarrow NS \nRightarrow CS$, core stability and Nash stability are independent of each other, i.e., a core stable partitions may not be Nash stable, and a Nash stable partition may not be core stable.

- $CS \Rightarrow PO \nRightarrow CS$, core stability implies Pareto optimality but the converse is not true.

- $IntS \nRightarrow NS \nRightarrow IntS$, internal stability and Nash stability are independent of each other.

- $IntS \nRightarrow PO \nRightarrow IntS$, internal stability and Pareto optimality are independent of each other.

- $NS \nRightarrow PO \nRightarrow NS$, Nash stability and Pareto optimality are independent of each other.

We note that for each hedonic game there exists an internally stable partition. Indeed, the partition that contains all singleton coalitions is internally stable. So, for each hedonic game $G = (N, \succeq)$ we have $IntS(N, \succeq) \neq \emptyset$. We also know that there exists a Pareto optimal partition for each hedonic game, i.e., for each $G = (N, \succeq)$ we have $PO(N, \succeq) \neq \emptyset$. These results together with the fact that internal stability and Pareto optimality are independent of each other give rise to the question that does there exist a partition which is both internally stable and Pareto optimal for each hedonic game. If the answer is yes, then the question is how can we find such a partition. We will show that for each hedonic game there exists a partition that is both internally stable and Pareto optimal. To do so, we will construct an algorithm that terminates for each hedonic game and produces a partition that is both internally stable and Pareto optimal.

3. AN ALGORITHM

In this section, we define an algorithm that finds an internally stable and Pareto optimal partition for a given hedonic game.

Let $\sigma: \{1,2,\ldots,n\} \to N$ be an ordering of agents, and $\Theta(N)$ denote the set of all orderings of agents. For an ordering σ, $\sigma(1)$ denote the first agent in the ordering, $\sigma(2)$ denote the second agent in the ordering, and so on. For a subset of agents $\tilde{N} \subseteq N$, we define the restriction of σ to $\tilde{N}$, $\sigma_{\tilde{N}}$, as follows: for each $i, j \in \tilde{N}$, $\sigma_{\tilde{N}}(i) < \sigma_{\tilde{N}}(j)$ if and only if $\sigma(i) < \sigma(j)$. For each $k \in \tilde{N}$, $\sigma_{\tilde{N}}(k)$ denotes the k^{th} agent in the ordering $\sigma_{\tilde{N}}$.

We now define a reduced game that will be used in the algorithm.

Definition 8 ***Reduced Game***

Let $G = (N, \succeq)$ be a hedonic game and $\tilde{N} \subseteq N$. The reduced game of $G = (N, \succeq)$ to $\tilde{N}$ equals $\tilde{G} = (\tilde{N}, \tilde{\succeq})$, where $\tilde{\succeq}$ is the restriction of $\succeq$ to $\tilde{N}$, $\tilde{\succeq} = \succeq_{\tilde{N}}$ i.e.,

(i) for each $i \in \tilde{N}$, $\tilde{\succeq}_i \in \mathcal{R}_i(\Sigma_i^{\tilde{N}})$ and

(ii) for each $S, T \in \Sigma_i^{\tilde{N}}$, $S \tilde{\succeq}_i T \Leftrightarrow S \succeq_i T$.

For each hedonic game $G = (N, \succeq)$ and an ordering of agents σ, we define the following algorithm:

Input. A hedonic game $G = (N, \succeq)$ and an ordering of agents $\sigma \in \Theta(N)$.

Step 1. Let $G_1 = G$, i.e., $N_1 := N$ and $\succeq_{N_1} := \succeq$. We consider the first agent $\sigma(1)$ in the ordering and her first-ranked (best) coalition $r_1(\sigma(1), \succeq_{\sigma(1)})$. If this coalition is internally stable, then we remove it (this coalition will be part of final partition).

If $r_1(\sigma(1), \succeq_{\sigma(1)})$ is not internally stable, we consider the second agent $\sigma(2)$ in the ordering and her best coalition $r_1(\sigma(2), \succeq_{\sigma(2)})$. If this coalition is internally stable, then we remove it (this coalition will be part of final partition).

If $r_1(\sigma(2), \succeq_{\sigma(2)})$ is not internally stable, we consider the third agent $\sigma(3)$ in the ordering and her best coalition. If this coalition is internally stable, then we remove it. If not, then we continue to choose agents according to the ordering σ and check the first-ranked coalitions of these agents. Whenever the first-ranked coalition of an agent is internally stable, we remove it.

If none of the best coalition of any agent is internally stable we choose agent $\sigma(1)$ again and consider now her second-ranked coalition $r_2(\sigma(1), \succeq_{\sigma(1)})$. If this coalition is internally stable, we remove it. If not, we choose agents according to σ and consider their second-ranked coalitions. That is, we first check first-ranked coalitions of all agents, and then second-ranked coalitions of all agents, and then third-ranked coalitions, and so on. In this process, when an internally stable coalition is seen, we remove it (it will be a part of the final partition).

Let H_1 denote the internally stable coalition that is removed at the end of Step 1. Let $N_2 = N_1 \backslash H_1$ be the set of remaining agents and, if $N_2 \neq \emptyset$, we continue with Step 2. Otherwise, we stop.

Step 2. We consider $G_2 = (N_2, \succeq_{N_2})$ and the ordering σ_{N_2}, i.e., we consider the reduced game with remaining set of agents N_2 and their restricted preference profile $\succeq_{N_2}$, and the restriction of σ to N_2, σ_{N_2}.

We consider the first agent $\sigma_{N_2}(1)$ and her first-ranked coalition $r_1(\sigma_{N_2}(1), \succeq_{\sigma_{N_2}(1)})$. If this coalition is internally stable, we remove it (this coalition will be part of the final partition). If it is not internally stable, then we consider the second agent $\sigma_{N_2}(2)$ and her first-ranked coalition. If this coalition is internally stable, then we remove it. Otherwise, we choose agents according to the ordering σ_{N_2} and check the first-ranked coalitions of these agents. Whenever the first-ranked coalition of an agent is internally stable, we remove it. If none of the best coalition of any agent in N_2 is internally stable we choose agent $\sigma_{N_2}(1)$ again and consider now her second-ranked coalition $r_2(\sigma_{N_2}(1), \succeq_{\sigma_{N_2}(1)})$. If this coalition is internally stable, we remove it. If not, we choose agents according to σ_{N_2} and consider their second-ranked coalitions. We first check the first-ranked coalitions of all agents in N_2, then second-ranked coalitions of all agents, then third-ranked coalitions, and so on. In this process, when an internally stable coalition is seen, we remove it (it will be a part of the final partition).

Let H_2 denote the internally stable coalition that is removed at the end of Step 2. Let $N_3 = N_2 \backslash H_2$ be the set of remaining agents and, if $N_3 \neq \emptyset$, we continue with Step 3. Otherwise, we stop.

In general, at Step t we have the following:

Step *t*. We consider $G_t = (N_t, \succeq_{N_t})$ and the ordering σ_{N_t}, i.e., we consider the reduced game with remaining set of agents $N_t = N \backslash \cup_{k=1}^{t-1} H_k$ after Step $t-1$ and their restricted preference profile $\succeq_{N_t}$, and the restriction of σ to N_t, σ_{N_t}.

We choose agents according to ordering σ_{N_t} and consider their first-ranked coalitions. Whenever the first-ranked coalition of an agent is internally stable we remove it. If none of the first-ranked coalitions of any agent in N_t is internally stable, then we consider the second-ranked coalitions of agents, and so on. When an internally stable coalition is detected in this process we remove it (it will be a part of the final partition). Let H_t denote the internally stable coalition that is removed at the end of Step t. Let $N_{t+1} = N_t \backslash H_t$ be the set of remaining agents and, if $N_{t+1} \neq \emptyset$, we continue with Step $t+1$. Otherwise, we stop.

Output. The algorithm terminates when the set of remaining agents is empty. When the algorithm terminates at the end of Step t^*, the output of the algorithm is the partition π^* that consists of the coalitions that are removed at the end of each step in the algorithm, i.e., $\pi^* = \{H_1, \dots, H_{t^*}\}$.

It is clear that the algorithm stops for each hedonic game and each ordering of agents.

We note that for a given hedonic game, we can find all internally stable and Pareto optimal partitions by running the algorithm for all orderings of agents.

Now we proceed with an example to show how the algorithm we introduce works and we apply our algorithm to the hedonic game below.

Example 1 Let $G = (N, \succeq)$ be a hedonic game with $N = \{1,2,3,4,5\}$ and the preference profile $\succeq = (\succeq_1, \succeq_2, \succeq_3, \succeq_4, \succeq_5) \in \mathcal{R}^N$ is given below:

$\succeq_1: \{1,2,3\} \succ_1 \{1,3\} \succ_1 \{1,5\} \succ_1 \{1,3,5\} \succ_1 \{1,4,5\} \succ_1 \{1\} \succ_1 \dots$,

$\succeq_2: \{1,2\} \succ_2 \{1,2,3\} \succ_2 \{2,4,5\} \succ_2 \{2,5\} \succ_2 \{2\} \succ_2 \dots$,

$\succeq_3: \{2,3\} \succ_3 \{3,4,5\} \succ_3 \{1,3\} \succ_3 \{1,3,5\} \succ_3 \{3\}, \succ_3 \dots$,

$\succeq_4: \{1,4,5\} \succ_4 \{1,2,4\} \succ_4 \{3,4,5\} \succ_4 \{4\} \succ_4 \dots$, and

$\succeq_5: \{1,3,5\} \succ_5 \{1,5\} \succ_5 \{3,4,5\} \succ_5 \{2,5\} \succ_5 \{1,4,5\} \succ_5 \{4,5\} \succ_5 \{5\} \succ_5 \dots$.

We now apply the algorithm to this hedonic game.

Let $\sigma: \{1,2, \dots, n\} \to N$ be an identity function, i.e., for each $i \in N$, $\sigma(i) = i$.

Step 1. Let $N_1 := N$ and $\succeq_{N_1} := \succeq$. We consider the hedonic game $G_1 = (N_1, \succeq_{N_1})$.

- We consider the first agent $\sigma(1) = 1$ in the ordering of σ. The first-ranked (best) coalition for agent *1* is $\{1,2,3\}$, i.e., $r_1(\sigma(1), \succeq_{\sigma(1)}) = r_1(1, \succeq_1) = \{1,2,3\}$. Coalition $\{1,2,3\}$ is not internally stable since coalition $\{3\}$ internally blocks it, i.e., $\{3\} \succ_3 \{1,2,3\}$.

- We consider the second agent $\sigma(2) = 2$ in the ordering of σ. The first-ranked (best) coalition for agent *2* is $\{1,2\}$, that is, $r_1(\sigma(2), \succeq_{\sigma(2)}) = r_1(2, \succeq_2) = \{1,2\}$. Coalition $\{1,2\}$ is not internally stable because coalition $\{1\}$ internally blocks it.

- We now consider the third agent $\sigma(3) = 3$, and $r_1(\sigma(3), \succeq_{\sigma(3)}) = r_1(3, \succeq_3) = \{2,3\}$. Coalition $\{2,3\}$ is not internally stable because coalition $\{2\}$ internally blocks it.

- We consider the fourth agent $\sigma(4) = 4$, and the best coalition for agent 4 is $\{1,4,5\}$ that is not internally stable since coalition $\{1,5\}$ internally blocks it.

- We consider the fifth agent $\sigma(5) = 5$, and the best coalition for agent 5 is $\{1,3,5\}$. Coalition $\{1,3,5\}$ is not internally stable since coalition $\{1,3\}$ internally blocks it.

None of the best coalition of any agent is internally stable. We now consider second-ranked coalitions of agents according to the ordering of σ. We choose agent $\sigma(1) = 1$ again and consider $r_2(\sigma(1), \succsim_{\sigma(1)}) = r_2(1, \succsim_1) = \{1,3\}$. Coalition $\{1,3\}$ is internally stable. We call $\{1,3\} = H_1$ and remove it. Let $N_2 = N_1 \backslash H_1 = \{2,4,5\}$. Since $N_2 \neq \emptyset$, we continue with Step 2.

Step 2. We consider $G_2 = (N_2, \succsim_{N_2})$, where $N_2 = N_1 \backslash H_1 = \{2,4,5\}$ and preferences $\succsim_{N_2} = (\tilde{\succsim}_2, \tilde{\succsim}_4, \tilde{\succsim}_5) \in \mathcal{R}^{N_2}$ are as follows:

$\tilde{\succsim}_2 : \{2,4,5\} \,\tilde{\succ}_2\, \{2,5\} \,\tilde{\succ}_2\, \{2\} \,\tilde{\succ}_2\, \ldots,$

$\tilde{\succsim}_4 : \{4\} \,\tilde{\succ}_4\, \ldots$, and

$\tilde{\succsim}_5 : \{2,5\} \,\tilde{\succ}_5\, \{4,5\} \,\tilde{\succ}_5\, \{5\} \,\tilde{\succ}_5\, \ldots.$

We consider the restriction of σ to N_2, σ_{N_2}. The first agent in the ordering of σ_{N_2} is agent *2*, i.e., $\sigma_{N_2}(1) = 2$. The first-ranked coalition of agent *2* is $\{2,4,5\}$. The coalition $\{2,4,5\}$ is not internally stable since coalition $\{4\}$ internally blocks it, i.e., $\{4\} \,\tilde{\succ}_4\, \{2,4,5\}$. We consider the second agent $\sigma_{N_2}(2) = 4$ and the first-ranked coalition for agent 4 is $\{4\}$ which is internally stable. We call $\{4\} = H_2$ and remove it. Let $N_3 = N_2 \backslash H_2 = \{2,5\}$. Since $N_3 \neq \emptyset$, we continue with Step 3.

Step 3. We consider $G_3 = (N_3, \succsim_{N_3})$, where $N_3 = N_2 \backslash H_2 = \{2,5\}$ and preferences $\succsim_{N_3} = (\hat{\succsim}_2, \hat{\succsim}_5) \in \mathcal{R}^{N_3}$ are as follows:

$\hat{\succsim}_2 : \{2,5\} \,\hat{\succ}_2\, \{2\} \,\hat{\succ}_2\, \ldots$, and

$\hat{\succsim}_5 : \{2,5\} \,\hat{\succ}_5\, \{5\} \,\hat{\succ}_5\, \ldots.$

We consider the restriction of σ to N_3, σ_{N_3}. The first agent in the ordering of σ_{N_3} is agent *2*, i.e., $\sigma_{N_3}(1) = 2$. The first-ranked coalition of agent *2* is $\{2,5\}$. The coalition $\{2,5\}$ is internally stable. We call $\{2,5\} = H_3$ and remove it. Let $N_4 = N_3 \backslash H_3 = \emptyset$. Since $N_4 = \emptyset$, we stop.

The output of the algorithm is $\pi^* = \{H_1, H_2, H_3\} = \{\{1,3\}, \{4\}, \{2,5\}\}$. The partition π^* is internally stable and Pareto optimal.

We note that the partition π^* is not core stable since the coalition $\{3,4,5\}$ blocks it, that is, for agent 3 we have $\{3,4,5\} \succ_3 \{1,3\} = \pi^*(3)$, for agent 4 we have $\{3,4,5\} \succ_4 \{4\} = \pi^*(4)$, and for agent 5 we have $\{3,4,5\} \succ_5 \{2,5\} = \pi^*(5)$.

We note that the partition π^* is not Nash stable since agent 2 Nash blocks π^*, i.e., agent 2 leaves her coalition $\pi^*(2) = \{2,5\}$ and joins coalition $\{1,2\} \in \pi^*$, and $\{1,2,3\} \succ_2 \{2,5\}$.

We now introduce our main result that shows the existence of an internally stable and Pareto optimal partition for each hedonic game.

Theorem 1 *For each hedonic coalition formation game $G = (N, \succeq)$ and each ordering $\sigma: \{1,2, \dots, n\} \to N$ of agents, the algorithm produces a partition that is both internally stable and Pareto optimal. In particular, $IntS(N, \succeq) \cap PO(N, \succeq) \neq \emptyset$.*

Proof. Let $G = (N, \succeq)$ be a hedonic game and σ be an ordering of agents. Suppose that the algorithm stops at Step t^* and $\pi^* = \{H_1, \dots, H_{t^*}\}$ is the partition obtained by the algorithm.

It is clear that every coalition in π^* is internally stable, otherwise they would not be chosen in the algorithm. So, the partition π^* is internally stable.

We now show that π^* is Pareto optimal. Suppose to the contrary that it is not Pareto optimal. Then, there exists another partition π that Pareto dominates π^*. Let $N' \subseteq N$ be the subset of agents such that their coalitions are different at partitions π and π^*, that is, $N' = \{i \in N | \pi^*(i) \neq \pi(i)\}$. Since agents have strict preferences and π Pareto dominates π^*, we have that for each $i \in N'$, $\pi \succ_i \pi^*$. This fact, together with π^* is internally stable, implies that the partition π is also internally stable.

Let $j \in N'$ be the agent such that for each $i \in (N' \backslash \{j\})$, $\sigma(j) < \sigma(i)$. We now consider the coalition $\pi^*(j)$. Note that for each $i \in \pi^*(j)$ we have $i \in N'$. The coalition $\pi^*(j)$ was removed at the end of some step of the algorithm and at that step agent j was chosen. However, since π is internally stable and for agent j we have $\pi(j) \succ_j \pi^*(j)$, the coalition $\pi(j)$ would have been removed in the algorithm and be a part of the partition π^*, a contradiction. Hence, π^* is Pareto optimal.

Therefore, $\pi^* \in IntS(N, \succeq) \cap PO(N, \succeq)$, and $IntS(N, \succeq) \cap PO(N, \succeq) \neq \emptyset$. □

Next, we explore the properties of the partition brought by the algorithm when the partition reflects no cooperation or full cooperation, i.e., we explore the properties of partitions that consists of singleton coalitions or the grand coalition.

We show that if the algorithm produces the partition at which every coalition consists of a singleton agent, then it is also core stable.

Proposition 1 *For any hedonic game $G = (N, \succeq)$ and any ordering $\sigma: \{1,2, \dots, n\} \to N$ of agents, if the algorithm produces the partition $\pi = \{\{1\}, \{2\}, \dots, \{n\}\}$, then π is core stable.*

Proof. Let $G = (N, \succeq)$ be a hedonic game and σ be an ordering of agents. Let $\pi = \{\{1\}, \{2\}, \dots, \{n\}\}$ be the outcome of the algorithm. By Theorem 1, π is internally stable and Pareto optimal.

We will show that π is core stable. Suppose not. Then, there exists a blocking coalition $H \subseteq N$ for partition π, i.e., for each $i \in H$ we have $H \succ_i \pi(i)$. We consider the partition π' that results in when coalition H blocks π, i.e., for each $i \in H$, $\pi'(i) = H$, and for each $j \in (N \backslash H)$, $\pi'(j) = \pi(j) = \{j\}$. So, $\pi' = \{H, \{\{j\} \big| j \in (N \backslash H)\}\}$. We now have that for each $i \in H$, $\pi' \succ_i \pi$, and for each $j \in (N \backslash H)$,

$\pi' \sim_j \pi$. This means that π' Pareto dominates π, which is a contradiction. Hence, $\pi = \{\{1\},\{2\},\dots,\{n\}\}$ is core stable. □

We show that if the algorithm produces the partition that contains only the grand coalition, then it is also both core stable and Nash stable.

Proposition 2 *For any hedonic game $G = (N, \succeq)$ and any ordering $\sigma: \{1,2,\dots,n\} \to N$ of agents, if the algorithm produces the partition $\pi = \{\{N\}\}$, then π is core stable and Nash stable.*

Proof. Let $G = (N, \succeq)$ be a hedonic game and σ be an ordering of agents. Let $\pi = \{\{N\}\}$ be the outcome of the algorithm. By Theorem 1, π is internally stable and Pareto optimal.

We first show that $\pi = \{\{N\}\}$ is core stable. Suppose not. Then, there exists a blocking coalition $H \subseteq N$ for π, i.e., for each $i \in H$, $H \succ_i \pi(i)$. However, this implies that coalition H also internally blocks π, since $\pi = \{\{N\}\}$. This is in contradiction with that π is internally stable. Therefore, π is core stable.

We now show that $\pi = \{\{N\}\}$ is Nash stable. Suppose not. Then, there exist an agent $i \in N$ and a coalition $H \in (\pi \cup \{\emptyset\})$ such that $H \cup \{i\} \succ_i \pi(i)$. However, since $\pi(i) = \{N\}$, this implies that $\{i\} \succ_i \pi(i)$, i.e., agent i also internally blocks π, a contradiction. Therefore, the partition π is Nash stable.□

4. CONCLUSION

We consider the model of hedonic coalition formation games that is used to model coalition formation problems without externalities. A hedonic coalition formation game consists of a finite set of agents and each agent has preferences over all coalitions containing her, i.e., each agent only cares about the members of her coalition and does not care how other agents form coalitions. Some examples are the production of local public goods, forming clubs, and so on. The outcome of a hedonic coalition formation game is a partition. Nash stability and core stability of partitions have been studied widely in the literature. Nash stability based on a single agent deviation and core stability based on a coalitional deviation from a given partition. Core stability also requires coordination of agents who are at different coalitions in the partition. When such a coordination is impossible then studying core stability is not an appropriate issue.

In this study, we focus on internal stability as a stability notion and Pareto optimality as an efficiency notion. Internal stability requires coordination of agents who are in the same coalition in a given partition. So, it is a weaker stability notion than core stability, i.e., core stability implies internal stability but internal stability does not imply core stability. We study the existence of internally stable and Pareto optimal partition for hedonic coalition formation games. We construct an algorithm such that it terminates in finite step for each hedonic game and show that the outcome of the algorithm is an internally stable and Pareto optimal partition. Moreover, we show that if the outcome of the algorithm is the partition that contains only singleton coalitions then it is also core stable. In addition, we show that if the outcome of the algorithm is the partition that contains only the grand coalition then it is also both Nash stable and core stable. A further study is to focus on the uniqueness of internally stable and Pareto optimal partition, that is, to determine conditions that guarantee a unique internally stable and Pareto optimal partition.

REFERENCES

Alcalde, J. and Revilla, P. (2004) "Researching with Whom? Stability and Manipulation", Journal of Mathematical Economics, 40: 869–887.

Alcalde, J. and Romero-Medina, A. (2006) "Coalition Formation and Stability", Social Choice and Welfare, 27: 365–375.

Aziz, H. and Brandl, F. (2012) "Existence of Stability in Hedonic Coalition Formation Games", Arxiv Preprint arXiv:1201.4754.

Aziz, H. and Savani, R. (2016) "Hedonic Games" F. Brandt, V. Conitzer, J. Lang U. Endriss, and AD Procaccia (eds.) Handbook of Computational Social Choice, Cambridge University Press, Cambridge.

Aziz, H., Brandt, F., and Harrenstein, P. (2013) "Pareto Optimality in Coalition Formation", Games and Economic Behavior, 82: 562–581.

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

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

Burani, N. and Zwicker, W. S. (2003) "Coalition Formation Games with Separable Preferences", Mathematical Social Sciences, 45: 27–52.

Cechlárová, K. and Romero-Medina, A. (2001) "Stability in Coalition Formation Games", International Journal of Game Theory, 29(4): 487–494.

Dimitrov, D. and Sung, S. C. (2004) "Enemies and Friends in Hedonic Games: Individual Deviations, Stability and Manipulation", CentER Discussion Paper Series.

Dimitrov, D. and Sung, S. C. (2006) "Top Responsiveness and Nash Stability in Coalition Formation Games", Kybernetika, 42(4): 453–460.

Dimitrov, D., Borm, P., Hendrickx, R., and Sung, S. C. (2006) "Simple Priorities and Core Stability in Hedonic Games", Social Choice and Welfare, 26(2): 421–433.

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

Gale, D. and Shapley, L. S. (1962) "College Admissions and the Stability of Marriage", American Mathematical Monthly, 69: 9–15.

Hajduková, J. (2006) "Coalition Formation Games: A Survey", International Game Theory Review, 8(4): 613–641.

Iehlé, V. (2007) "The Core-Partition of a Hedonic Game", Mathematical Social Sciences, 54: 176–185.

İnal, H. (2019) "The Existence of a Unique Core Partition in Coalition Formation Games", Games and Economic Behavior, 114: 215–231.

Karakaya, M. (2011) "Hedonic Coalition Formation Games: A New Stability Notion", Mathematical Social Sciences, 61: 157–165.

Liu, Y., Tang, P., and Fang, W. (2014) "Internally Stable Matchings and Exchanges", In Proceedings of the AAAI Conference on Artificial Intelligence, 28: 1433–1439.

Özbilen, S. (2019) "Three Essays on Coalition Formation Games", Ph.D. Thesis, İstanbul Bilgi University.

Özbilen, S. (2022) "Coalition Formation Games: Internal Stability", forthcoming in Ankara University SBF Journal.

Pápai, S. (2004) "Unique Stability in Simple Coalition Formation Games", Games and Economic Behavior, 48: 337–354.

Pápai, S. (2007) "Individual Stability in Hedonic Coalition Formation", Mimeo.

Roth, A. E. and Sotomayor, M. A. O. (1990) "Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis", Cambridge University Press, Cambridge.

Schlueter, J. and Goldsmith, J. (2020) "Internal Stability in Hedonic Games", In The Thirty-Third International Flairs Conference: 154–159.

Suksompong, W. (2015) "Individual and Group Stability in Neutral Restrictions of Hedonic Games", Mathematical Social Sciences, 78: 1–5.

Sung, S. C. and Dimitrov, D. (2007) "On Myopic Stability Concepts for Hedonic Games", Theory and Decision, 62: 31–45.

Suzuki, K. and Sung, S. C. (2010) "Hedonic Coalition Formation in Conservative Societies", SSRN: 1700921.

KATKI ORANI/ CONTRIBUTION RATE AÇIKLAMA	AÇIKLAMA / EXPLANATION	KATKIDA BULUNANLAR / CONTRIBUTORS
Fikir veya Kavram / Idea or Notion	Araştırma hipotezini veya fikrini oluşturmak / Form the research hypothesis or idea	Asst. Prof. (Ph.D.) Mehmet KARAKAYA Asst. Prof. (Ph.D.) Seçkin ÖZBİLEN
Tasarım / Design	Yöntemi, ölçeği ve deseni tasarlamak / Designing method, scale and pattern	Asst. Prof. (Ph.D.) Mehmet KARAKAYA Asst. Prof. (Ph.D.) Seçkin ÖZBİLEN
Veri Toplama ve İşleme / Data Collecting and Processing	Verileri toplamak, düzenlenmek ve raporlamak / Collecting, organizing and reporting data	Asst. Prof. (Ph.D.) Mehmet KARAKAYA Asst. Prof. (Ph.D.) Seçkin ÖZBİLEN
Tartışma ve Yorum / Discussion and Interpretation	Bulguların değerlendirilmesinde ve sonuçlandırılmasında sorumluluk almak / Taking responsibility in evaluating and finalizing the findings	Asst. Prof. (Ph.D.) Mehmet KARAKAYA Asst. Prof. (Ph.D.) Seçkin ÖZBİLEN
Literatür Taraması / Literature Review	Çalışma için gerekli literatürü taramak / Review the literature required for the study	Asst. Prof. (Ph.D.) Mehmet KARAKAYA Asst. Prof. (Ph.D.) Seçkin ÖZBİLEN

Hakem Değerlendirmesi: Dış bağımsız.

Çıkar Çatışması: Yazar çıkar çatışması bildirmemiştir.

Finansal Destek: Yazar bu çalışma için finansal destek almadığını beyan etmiştir.

Teşekkür: -

Peer-review: Externally peer-reviewed.

Conflict of Interest: The author has no conflict of interest to declare.

Grant Support: The author declared that this study has received no financial support.

Acknowledgement: -