Economics & Management Series EMS-2021-05

# *Marginal Productivity and Coalition Formation with Distributive Norms*

Hideaki Goto
*International University of Japan*

November 2021

IUJ Research Institute
International University of Japan

# Marginal Productivity and Coalition Formation with Distributive Norms

Hideaki Goto*

**Abstract**

This paper analyzes coalition formation under constant, decreasing, and increasing marginal productivity when the total surplus jointly produced by individuals with heterogeneous abilities can only be distributed to its members in egalitarian or meritocratic ways. When marginal productivity is decreasing or constant, the results are simple, as no coalition with multiple members is included in a stable coalition structure when marginal productivity is decreasing, whereas individuals are indifferent to which *meritocratic* coalition they belong, including singletons, in the case of constant marginal productivity. In contrast, if marginal productivity is increasing, stable structures differ considerably from those obtained by other models. A procedure to identify stable structures is proposed, finding that multiple egalitarian coalitions can exist, each of which is always consecutive, but there is, at most, only one meritocratic coalition, which may or may not be consecutive, in stable structures. Moreover, the grand egalitarian coalition is only stable under certain conditions, whereas the grand meritocratic coalition is always stable.

*JEL Classification:* C71, D71, D63, D30.

*Keywords*: Coalition formation; Egalitarianism; Meritocracy; Marginal productivity.

*International University of Japan, 777 Kokusai-cho, Minami Uonuma-shi, Niigata 949-7277, Japan. Email: h-goto@iuj.ac.jp.

# 1 Introduction

Many human activities are performed by groups of people, ranging from firms to political parties to friend groups. To attain higher utility (or *payoff*), individuals either decide which group (or coalition) to belong to or form a new coalition independently. If all the people in each existing coalition are satisfied, the coalitions are *stable*, whereas if another coalition gives a higher payoff, they move to that coalition or form such a new coalition. What coalition structures are stable under such circumstances? This problem has been actively studied in the field of *coalition formation* (Ray, 2007; Ray and Vohra, 2014).

This paper aims to examine coalition formation in the following situations: A society consists of $n$ individuals with heterogeneous abilities. The total surplus of a coalition rises by total ability, *not* the total number of individuals, in the coalition. There is no minimum size for a coalition to be deemed productive; that is, each individual can receive payoff according to a production function without teaming up with other individuals. Marginal productivity can be either decreasing, constant, or increasing. Because of social norms and/or the considerable costs associated with more complex payoff calculations, the total surplus of a coalition can be distributed among its members in either an egalitarian or meritocratic way only. A coalition structure, or more simply, a *structure*, is *stable* if there is no other coalition all the members of which receive strictly higher payoffs than in the structure. What structures are stable under differing marginal productivities in such situations? What are the characteristics of coalitions in such stable structures? How can stable structures be identified? To the best of our knowledge, this paper is the first to address these questions in this context.

As is easily expected, no coalition (with multiple members) is formed when marginal productivity is decreasing. When marginal productivity is constant, stable structures do not contain any egalitarian coalition because no individual will equally share total surplus with lower-ability individuals. If total surplus is allocated in proportion to ability, payoffs remain unchanged, whether or not individuals form a coalition. Thus, any structures with meritocratic coalitions, including singletons, are stable in cases of constant marginal productivity.

The analysis becomes more complex when marginal productivity is increasing, wherein adding a lower-ability individual to a coalition may increase the egalitarian payoff to its members. In the special case in which average total surplus is maximized in a grand coalition, the egalitarian grand coalition is stable. More generally, meritocratic payoffs always increase as the coalition expands, whereas egalitarian payoffs may or may not increase as a new member is added. Furthermore, it is possible that the payoff of a member of an egalitarian coalition increases by joining a meritocratic coalition, even if the average total surplus is maximized by the egalitarian coalition. A procedure to identify stable structures is then proposed, demonstrating that egalitarian and meritocratic coalitions may coexist in a stable structure. Furthermore, in stable structures, multiple egalitarian coalitions can exist, each of which is always consecutive, whereas there is, at most, one meritocratic coalition, which may or may not be consecutive. In any case, grand meritocratic coalitions are always stable when marginal productivity is increasing.[1]

We consider a hedonic game (Banerjee et al., 2001; Bogomolnaia and Jackson, 2002), in which there are no externalities across coalitions so that each individual's payoff is completely determined by the members of a coalition to which (s)he belongs. This paper is broadly related to the literature on coalition formation by individuals with multi-dimensional preferences or preferences depending on non-material payoff. Watts (2007) analyzes two types of agents separately, one that seeks to be the highest status agent within the group and another who wants to be a member of the highest status group that (s)he can join. Razin and Piccione (2009) study partition function games in which each agent's social ranking is determined by power relations between coalitions as well as the agent's ranking within the coalition to which (s)he belongs, and each agent prefers to be ranked higher in the society. Morelli and Park (2016) consider the case in which the total surplus of a coalition is a function of the number of coalition members (or *size*) and their aggregate ability (or *power*). For a coalition to produce positive total surplus, its power must exceed a certain minimum threshold. A coalition is *efficient* if both its size and power exceed the corresponding thresholds, and the total surplus of an efficient coalition is simply its power. Among other results, they demonstrate that there is no profitable coalitional deviation from a coalition partition if and only if each coalition is efficient and each member's payoff is equal to her/his ability.

Distribution rules have been extensively studied (see, for example, Moulin (1987) and Roemer and Silvestre (1993) for representative works in this field. See also Moulin (1988, 2003) and Roemer (1996) for book-length treatments of this and related topics). Among them, the most frequently considered are egalitarian and proportional sharing rules, which are also incorporated into the analyses of coalition formation as distributive norms. Farrell and Scotchmer (1988) is a pioneering work that studies core partitions when the total surplus is equally shared within each coalition. More recently, Herings et al. (2021) analyze two types of societies *separately*: one in which the total surplus of each coalition is split equally and the other in which the total surplus is split according to individual productivity. In their model, individuals care about their relative payoffs within the coalition as well as material payoffs.

[1]Therefore, the existence of stable structures, which is studied by Banerjee et al. (2001), for example, is not the main concern of the current research.

Coalitions produce a positive total surplus when their size reaches a certain threshold. If the size is large enough, the total surplus is the sum of the members' productivities. They fully characterize core partitions in each society. The main difference between previous studies and the current research is that they consider societies with only one distributive norm, whereas we consider the case in which two distribution rules exist.

More closely related to this study is Barberà et al. (2015), who propose a model in which coalition members vote between egalitarian and meritocratic distributions of joint benefits. In their main analysis, people have one of three productivities (low, medium, and high), or can be clustered into three classes according to productivities. Coalitions smaller than a certain size produce nothing, whereas larger coalitions produce the sum of the members' productivities. The conditions for the existence of core stable coalition structures are identified, also demonstrating that meritocratic and egalitarian coalitions can coexist in a stable coalition structure. However, in contrast to the current study, the results are closely related to the minimum threshold size and its size relative to the numbers of individuals with different productivities. Moreover, the results are derived through a process that is different from ours: namely, voting. This study analyzes core partitions in a simple coalition formation game, assuming a simpler and more general production function with no minimum threshold size for coalitions to be productive. Furthermore, arbitrarily many levels of individual abilities are considered, mainly focusing on the case of increasing marginal productivity. To our knowledge, a simple coalition formation game with two distribution rules to choose from has not been analyzed in the literature. This work seeks to fill this gap.[2]

The remainder of this paper is organized as follows. Section 2 describes the model. Section 3 presents our results. Section 4 concludes the paper and discusses the issues that remain unsolved.

# 2 Model

We consider a set of $n$ individuals, $N = \{1, 2, \dots, n\}$, with heterogeneous abilities, $a = \{a_1, \dots, a_n\}$. Without loss of generality, the individuals are indexed in a decreasing order of ability: $a_1 > a_2 > \dots > a_n > 0$. A *coalition* $S$ is a non-empty subset of $N$, $S \subseteq N$, where $N$ is called the *grand coalition*. A *subcoalition* $S'$ of a coalition $S$ is a subset of $S$, $S' \subseteq S$. Proper inclusion is denoted by $\subset$. Production takes place through a coalition. The

[2]Additionally, although people often form coalitions to achieve higher marginal productivity, such cases have rarely been the main focus of past studies. Exceptions include Farrell and Scotchmer (1988), but as already noted, they consider a society in which only one distribution rule, an egalitarian division, is used.

total surplus of a coalition $S$ is given by $f\left(\sum_{i\in S} a_i\right)$, where $f$ is a strictly increasing, twice differentiable function with $f(0)=0$. A *society* is represented by $(N,a,f)$.

We assume that social norms dictate how a coalition distributes its total surplus to its members, and it can only choose between two division rules, that is, *egalitarian* division and *meritocratic* division. The former divides the total surplus equally among members, whereas the latter distributes the total surplus proportional to the members' abilities. Let $u_i^e$ and $u_i^m$ represent the payoffs to individual $i$ when the total surplus is distributed according to the egalitarian and meritocratic divisions, respectively. Then we have:

$$\text{Egalitarian division:}\quad u_i^e(S)=\frac{f\left(\sum_{i\in S} a_i\right)}{|S|}=\bar{a}_S\bar{f}\left(\sum_{i\in S} a_i\right)\quad \text{for all } i\in S \tag{1}$$

$$\text{Meritocratic division:}\quad u_i^m(S)=\frac{a_i}{\sum_{i\in S} a_i}f\left(\sum_{i\in S} a_i\right)=a_i\bar{f}\left(\sum_{i\in S} a_i\right)\quad \text{for } i\in S, \tag{2}$$

where $|S|$ is the number of members in $S$, $\bar{a}_S$ is the average ability of the members in $S$, $\bar{a}_S=\left(\sum_{i\in S} a_i\right)/|S|$, and $\bar{f}$ is the average total surplus per *ability*, $\bar{f}\left(\sum_{i\in S} a_i\right)=f\left(\sum_{i\in S} a_i\right)/\left(\sum_{i\in S} a_i\right)$. A coalition that employs the egalitarian (meritocratic) division rule will be called an *egalitarian coalition* (*meritocratic coalition*). The payoff of an individual when (s)he belongs to an egalitarian (meritocratic) coalition is her/his *egalitarian (meritocratic) payoff*.

A *structure* is a pair $(\pi,\rho)$, where $\pi=\{S_1,\ldots,S_H\}$ is a partition of $N$ and $\rho$ is a function that specifies a division rule of each coalition. A coalition *blocks* a structure if the payoff of each member of the coalition is strictly higher than the payoffs allocated in the structure. A structure is *stable* if no coalition blocks it.

# 3 Stable structures

This section presents our main results. We consider stable structures for the cases of constant, decreasing, and increasing marginal productivity.

## 3.1 Constant marginal productivity

We will first consider linear production functions. In this case, if a higher-ability individual belongs to the same egalitarian coalition as lower-ability individuals, her/his payoff is lower than when (s)he is alone. Since the individuals' abilities are assumed to be heterogeneous, no egalitarian coalition exists in a stable structure. Furthermore, as the average total surplus per ability is constant, as per equation (2), the meritocratic payoff of each individual does not depend on which meritocratic coalition they belong to. Therefore, the following proposition holds for the case of constant marginal productivity:

**Proposition 1.** *If $f$ is linear ($f'' = 0$), then any partition with meritocratic division in each coalition, including singletons and the grand coalition, is stable.*

*Proof.* If there is an egalitarian coalition with multiple members, then the payoff of (at least) the highest-ability individual increases through exiting the coalition alone and forming a singleton; thus, there is no egalitarian coalition (with more than one member) in a stable structure. Since the average total surplus per ability, $\bar{f}$, is constant, each individual's meritocratic payoff remains unchanged, regardless of what coalition (s)he belongs to, including a singleton. Thus, no meritocratic coalition blocks a structure consisting of only meritocratic coalitions. It is obvious that no egalitarian coalition blocks such partitions. $\square$

As a result, each individual $i \in N$ receives $f(a_i)$ in stable structures when the production function is linear.

## 3.2 Decreasing marginal productivity

A few basic facts will be used repeatedly to obtain results for cases of decreasing and increasing marginal productivity, so we state those as lemmas. Let a set of consecutively indexed individuals be $S_{j,k} = \{j, j+1, \dots, k-1, k\}$ ($j \leq k$). In the special case of $j = k$, we define as $S_{j,j} = \{j\}$. For $j < k$, such coalitions are called *consecutive coalitions*, following Greenberg and Weber (1986) and Herings et al. (2021). The egalitarian payoff of individual $i \in S_{j,k}$ is given by $u_i^e(S_{j,k}) = f(\sum_{i=j}^{k} a_i)/(k - j + 1)$.

**Lemma 1.** *For $i \in S_{j,k}$, $u_i^e(S_{j,k})$ is decreasing in $k$ if $f'' \leq 0$, decreasing in $j$ if $f'' \geq 0$.*

*Proof.* $u_i^e(S_{j,k})$ is rewritten as

$$u_i^e(S_{j,k}) = \frac{\sum_{l \in S_{j,k}} \left[ f\left(\sum_{i=j}^{l} a_i\right) - f\left(\sum_{i=j}^{l-1} a_i\right)\right]}{|S_{j,k}|},$$

where $f(\sum_{i=j}^{j-1} a_i)$ is defined as 0.

Since $a_{k+1} < a_i$ for all $i \in S_{j,k}$, if $f'' \leq 0$, then $f(\sum_{i=j}^{k+1} a_i) - f(\sum_{i=j}^{k} a_i) < f(\sum_{i=j}^{l} a_i) - f(\sum_{i=j}^{l-1} a_i)$ for all $l \in S_{j,k}$; thus, adding individual $k+1$ to $S_{j,k}$ reduces the egalitarian payoff.

Suppose $f'' \geq 0$. Adding individual $j$ to $S_{j+1,k}$ increases total surplus by $f(\sum_{i=j}^{k} a_i) - f(\sum_{i=j+1}^{k} a_i)$. Since $a_j > a_i$ for all $i \in S_{j+1,k}$, we have $f(\sum_{i=j}^{k} a_i) - f(\sum_{i=j+1}^{k} a_i) > f(\sum_{i=j+1}^{l} a_i) - f(\sum_{i=j+1}^{l-1} a_i)$ for all $l \in S_{j+1,k}$. Therefore, the egalitarian payoff increases by adding individual $j$, that is, it is decreasing in $j$. $\square$

As the following example shows, $u_i^e(S_{j,k})$ may not be decreasing in $j$ (in $k$) if $f'' < 0$ ($f'' > 0$).

**Example 1.** Suppose $N = \{1, 2, 3\}$ and $a = \{5, 4, 3\}$. If $f(x) = x^{1/2}$, then $u_i^e(S_{2,3}) = \sqrt{7}/2 = 1.32... > 1.15... = \sqrt{12}/3 = u_i^e(S_{1,3})$. So $u_i^e(S_{j,3})$ is not decreasing in $j$. If $f(x) = x^2$, then $u_i^e(S_{1,2}) = 9^2/2 = 40.5 < 48 = 12^2/3 = u_i^e(S_{1,3})$; thus, $u_i^e(S_{1,k})$ is not decreasing in $k$.

The next lemma concerns the relationship between meritocratic payoff and coalition size for any, not necessarily consecutive, coalitions:

**Lemma 2.** *For $i \in S$, $u_i^m(T) > (<, =) u_i^m(S)$ for $S \subset T$ if $f'' > 0$ ($f'' < 0$, $f'' = 0$).*

*Proof.* $u_i^m(S)$ is given by the individual's ability $a_i$ multiplied by the average total surplus per ability $f\left(\sum_{i \in S} a_i\right) / \left(\sum_{i \in S} a_i\right)$, where the latter is strictly increasing (strictly deceasing, constant) in aggregate ability if $f'' > 0$ ($f'' < 0$, $f'' = 0$). Since the aggregate ability $\sum_{i \in S} a_i$ increases as $S$ expands, the lemma is proved. □

Using Lemmas 1 and 2, we obtain the following result for the case of decreasing marginal productivity.

**Proposition 2.** *If $f'' < 0$, then no individual forms a coalition with any other individual in stable structures.*

*Proof.* From the proof of Lemma 1, an individual's egalitarian payoff decreases if (s)he forms a coalition with lower-ability individuals. Based on Lemma 2, meritocratic payoff decreases if an individual forms a coalition with any other individual(s). Thus, no coalition with more than one member is formed if $f'' < 0$, and each individual $i \in N$ receives $f(a_i)$. □

## 3.3 Increasing marginal productivity

When marginal productivity is increasing, $f'' > 0$, the average total surplus per *ability* always increases as a coalition expands. Therefore, based on equation (2), an individual's meritocratic payoff increases as new members join the coalition to which (s)he belongs. Obviously, the egalitarian payoff increases as a higher-ability individual than the current members joins a coalition.[3] Depending on the degree of increase in marginal productivity and the ability of a newly added individual, the egalitarian payoff may also increase as a lower-ability individual joins the coalition. As shown in equation (1), the payoffs of those with higher-than-average (lower-than-average) abilities are lower (higher) than those of a meritocratic coalition with

[3] Lemma 1 shows that this is the case for consecutive coalitions. It will be shown in Lemma 4 that if stable structures contain egalitarian coalitions, then they are all consecutive.

the same aggregate ability. These factors can influence one another to determine a stable structure in the case of increasing marginal productivity.

First, note the following basic property of a meritocratic coalition and egalitarian subcoalitions.

**Lemma 3.** *Suppose $f'' \geq 0$ and consider a meritocratic coalition $S \subseteq N$. There is no egalitarian subcoalition $S' \subseteq S$ in which all its members receive a strictly higher payoff than in the meritocratic coalition $S$.*

*Proof.* Since $S' \subseteq S$, by Lemma 2, the meritocratic payoffs of all the individuals in $S'$ are (weakly) lower than those received in $S$. The payoff of at least the highest ability individual in $S'$ further decreases when $S'$ is egalitarian. □

Note that the above lemma holds for any egalitarian subcoalitions, whether consecutive or not. If we apply Lemma 3 to the grand meritocratic coalition, we obtain the following result:[4]

**Proposition 3.** *If $f'' > 0$, then the grand meritocratic coalition is stable.*

*Proof.* From Lemmas 2 and 3, no egalitarian or meritocratic coalition, whether consecutive or not, blocks the grand meritocratic coalition. □

If the degree of increase in marginal productivity is high and/or the distribution of abilities is sufficiently equal, the average total surplus may be maximized by the grand coalition. In such cases, the following proposition holds:

**Proposition 4.** *If $f'' > 0$ and if $u_i^e(S_{1,k})$ is maximized at $k = n$, then the grand egalitarian coalition is stable.*

*Proof.* (a) *No egalitarian coalition blocks the grand egalitarian coalition*: The higher the abilities of a coalition's members, the higher the average total surplus of the coalition. Thus, the coalition that attains the maximum average total surplus must include individual 1 and be consecutive; that is, $S_{1,k}$ for some $k \in N$. Since the average total surplus is maximized by the grand coalition $S_{1,n} \equiv N$, no egalitarian coalition blocks the grand egalitarian coalition.

(b) *No meritocratic coalition blocks the grand egalitarian coalition*: Suppose not. Then, a meritocratic coalition $S \subseteq N$ exists in which all members receive a strictly higher payoff than $u_i^e(N)$. This means that the average total surplus of $S$ is greater than that of the grand coalition, a contradiction to (a).

□

[4] As Proposition 1 shows, this result holds for $f'' = 0$ as well.

A special case in which the average total surplus is maximized by the grand coalition is when the egalitarian payoff of a consecutive coalition always increases as it allows individuals with lower abilities to join. Thus, the next corollary follows immediately from Proposition 4. Note that $u_i^e(S_{1,k})$ is strictly increasing in $k$ only when $f'' > 0$.

**Corollary 1.** *If $u_i^e(S_{1,k})$ is strictly increasing in $k \in N$, then the grand egalitarian coalition is stable.*

Needless to say, the average total surplus may not be maximized by the grand coalition. In particular, if the rate of increase in marginal productivity is not very high or if ability distribution is unequal, the average total surplus of a coalition would decrease as a lower-ability individual joins the coalition. We will then use the following feature to simplify the study of such cases.

**Lemma 4.** *If a stable structure contains any egalitarian coalitions, then they are all consecutive.*

*Proof.* From Propositions 1 and 2, a stable structure contains egalitarian coalitions only if $f'' > 0$. Suppose that, in a stable structure, there is a non-consecutive egalitarian coalition $S$ that includes $S'$ and $S''$, where $\max\{i \mid i \in S'\} + 1 < \min\{j \mid j \in S''\}$. Let individual $k$ be the lowest-ability individual in $S'$; that is, $k = \max\{i \mid i \in S'\}$. Suppose that individual $k+1$ belongs to coalition $T(\neq S)$, which is either egalitarian or meritocratic.

(a) If $u_k^e(S) \geq u_{k+1}^e(T)$ or $u_k^e(S) \geq u_{k+1}^m(T)$, then the egalitarian coalition $(S\backslash\{j\})\cup\{k+1\}$ with $j \in S''$ blocks the structure.

(b) If $u_k^e(S) < u_{k+1}^e(T)$, then the egalitarian coalition $(T\backslash\{k+1\})\cup\{k\}$ blocks the structure.

(c) If $u_k^e(S) < u_{k+1}^m(T)$, then the meritocratic coalition $T \cup \{k\}$ blocks the structure.

Since (a)–(c) exhaust all possibilities, there exists no non-consecutive egalitarian coalition in any stable structures. □

It then turns out that if adding a lower-ability individual to consecutive egalitarian coalitions always lowers the average total surplus, the grand meritocratic coalition is the only stable structure.

**Proposition 5.** *If $f'' > 0$ and $u_i^e(S_{j,k}) > u_i^e(S_{j,k+1})$ for all $j < k \leq n-1$, then the only stable structure is the grand meritocratic coalition.*

*Proof.* From Proposition 3, the grand meritocratic coalition is stable. Suppose that another stable structure exists. If that structure contains an egalitarian coalition, then from Lemma 4, it must be consecutive. However, from $u_i^e(S_{j,k}) > u_i^e(S_{j,k+1})$ for all $j < k \leq n-1$, the structure is blocked by an egalitarian coalition consisting of all members of the egalitarian coalition but the lowest-ability one. Thus, the stable structure must consist of only meritocratic coalitions. However, from Lemma 2, any structure with multiple meritocratic coalitions is blocked by the grand meritocratic coalition. Therefore, the proposition is proved. □

More generally, the egalitarian payoff may or may not decrease when a lower-ability individual joins the coalition. If the average total surplus decreases, the members of egalitarian coalitions will not allow lower-ability individuals to join. Moreover, egalitarian payoffs must be high enough to motivate higher-ability individuals to stay in the coalition. Combining these observations with Lemma 4, egalitarian coalitions in a stable structure, if any, must be consecutive, $S_{j,k}$ ($j < k$), such that the egalitarian payoff $u_i^e(S_{j,k})$ is maximized by $k$ given $j$. This corresponds to the *Choice of the Strongest Procedure* in Herings et al. (2021) (see also Farrell and Scotchmer, 1988). In this current model, however, individuals can choose to form not only an egalitarian coalition but also a meritocratic coalition; thus, their procedure helps find only *candidates* of egalitarian coalitions in a stable structure.

To seek stable structures for more general cases in the current setting, we need to extend their procedures to incorporate the possibility of the formation of a meritocratic coalition. As noted above, it is necessary for the average total surplus of an egalitarian coalition to be maximized, given the highest-ability member. Those who do not benefit from joining an egalitarian coalition will join a meritocratic one (or will not join any coalition). From Lemma 2, the members of meritocratic coalitions always benefit by accepting new members. Therefore, egalitarian payoffs must be high enough to keep members of egalitarian coalitions—especially higher-ability ones—from moving to or forming a meritocratic coalition.

Taking these points into consideration, let us consider the following procedure.

**Step 1.** Find $k \in N$ at which $u_i^e(S_{1,k})$ ($i \in S_{1,k}$) is maximized. If there is more than one maximizer, choose any one of them.

**Step 2.**

Case 1: If the maximizer in Step 1 is $k = 1$, then include the highest-ability individual in a meritocratic coalition $S^m$.

Case 2: If the maximizer in Step 1 is $k \neq 1$, then denote it by $k_1$ and form an egalitarian coalition $S_1 = \{1, \ldots, k_1\}$. If $k_1 = n$, the grand egalitarian coalition is formed and the procedure ends.

**Step 3.** If the grand coalition is not formed in Step 2, then apply Steps 1 and 2 to $N\backslash\{1\}$ in Case 1 and $N\backslash S_1$ in Case 2, instead of $N$. If an egalitarian coalition is formed in the former case (the latter case), denote it by $S_1$ ($S_2$).

**Step 4.** Repeat the above steps until every individual belongs to some newly formed coalition, which may also be singletons.

As a result of the above steps, one of the following two cases occurs:

**Case A.** *All individuals belong to the grand coalition*: The procedure ends in this case.

**Case B.** *There is at least one (non-grand) egalitarian coalition*: Let $(\pi, \rho)$ denote the structure. Let $S_1, \ldots, S_H$ ($H \geq 1$) be the egalitarian coalitions and $S'_1, \ldots, S'_H$ their subcoalitions, respectively. If there are no $S'_1, \ldots, S'_H$ such that the joint meritocratic coalition $S'_1 \cup \ldots \cup S'_H \cup S^m$ blocks $(\pi, \rho)$, where $S^m$ may be empty, then the procedure ends. Otherwise, proceed to Step 5.

**Step 5.** Add the largest set $S'_1 \cup \ldots \cup S'_H$ that satisfies the following condition to the meritocratic coalition $S^m$: all individuals in the set are strictly better off in the expanded meritocratic coalition, that is, $u_i^e(S_h) < u_i^m(S'_1 \cup \ldots \cup S'_H \cup S^m)$ for all $i \in S_h$ ($h = 1, \ldots, H$).[5] Let the subsets $S'_1, \ldots, S'_H$ in this largest set and the expanded meritocratic coalition be respectively denoted by $S''_1, \ldots, S''_H$ and $S^M$, and proceed to Step 6.

**Step 6.** Let $T^m \subset S^M$ be the set with the greatest aggregate ability such that all individuals $i \in T^m$ are strictly better off joining egalitarian coalitions that are formed by applying Steps 1–4 to $(S_1 \backslash S''_1) \cup \ldots \cup (S_H \backslash S''_H) \cup T^m$ than staying in the meritocratic coalition $(S^M \backslash T^m) \cup \{i\}$. (Note that $T^m$ may be empty.)[6]

Let $(\pi^*, \rho^*)$ be a structure resulting from the above procedure. The following proposition holds:

**Proposition 6.** *$(\pi^*, \rho^*)$ is stable.*

*Proof.* From Propositions 3 and 4, $(\pi^*, \rho^*)$ is stable when $\pi^*$ is the grand coalition, whether egalitarian or meritocratic. So suppose there is at least one (non-grand) egalitarian coalition in $\pi^*$.

[5]Individuals who belong to the same egalitarian coalition receive the same payoff. From (2), an individual with a higher ability receives a higher payoff in the meritocratic coalition. Therefore, if $S'_h$ is not empty, then it must include the highest ability individual in $S_h$ and be consecutive.

[6]If in Step 5 individuals in egalitarian coalitions *anticipate* that Steps 1–4 will be applied to those who remain in egalitarian coalitions and only those who are still better off moving to the meritocratic coalition join it, then Step 6 can be omitted.

From Step 6, those who benefit from joining egalitarian coalitions that accept them have already done so. Moreover, because Steps 1–4 have been applied to all individuals in egalitarian coalitions, $(S_1 \backslash S_1'') \cup \ldots \cup (S_H \backslash S_H'') \cup T^m$, there is no egalitarian coalition in which a subset of them can receive a strictly higher payoff. Hence, no egalitarian coalition blocks $(\pi^*, \rho^*)$.

Individuals in $(S_1 \backslash S_1'') \cup \ldots \cup (S_H \backslash S_H'')$ remained in egalitarian coalitions when they could move to $S^M$. As individuals in $T^m$ join egalitarian coalitions and Steps 1–4 are applied to $(S_1 \backslash S_1'') \cup \ldots \cup (S_H \backslash S_H'') \cup T^m$, the payoffs of those in $(S_1 \backslash S_1'') \cup \ldots \cup (S_H \backslash S_H'')$ increase (at least weakly), whereas the payoffs they can receive by joining the meritocratic coalition $S^M \backslash T^m$ decrease. Hence, their payoffs do not increase by joining the meritocratic coalition $S^M \backslash T^m$. From Step 6, individuals in $T^m$ also receive higher payoffs in egalitarian coalitions than by joining the meritocratic coalition $S^M \backslash T^m$. Thus, there is no meritocratic coalition that blocks $(\pi^*, \rho^*)$ either. Therefore, $(\pi^*, \rho^*)$ is stable. $\square$

As reflected in the above procedure and Proposition 6, since meritocratic payoffs increase as the coalition expands, the members of meritocratic coalitions are always willing to merge with other meritocratic coalitions, whether or not there are egalitarian coalitions. This naturally leads to the following result:

**Corollary 2.** *At most, there is one meritocratic coalition in a stable structure, which may not be consecutive.*

The following examples show how the above procedure is applied to simple societies.

**Example 2.** Suppose that $f(x) = x^2$, $N = \{1, 2, 3, 4\}$ and $a = \{10, 9, 2, 0.5\}$. From Steps 1–4, an egalitarian coalition $S_1^e = \{1, 2\}$ and a meritocratic coalition $S^m = \{3, 4\}$ are formed, where $u_1^e(S_1^e) = u_2^e(S_1^e) = 180.5$, $u_3^m(S^m) = 5$ and $u_4^m(S^m) = 1.25$. By applying Step 5, the grand meritocratic coalition is formed, in which the individuals' payoffs are $u_1^m(N) = 215$, $u_2^m(N) = 193.5$, $u_3^m(N) = 43$, and $u_4^m(N) = 10.75$.

**Example 3.** Suppose that $f(x) = x^2$, $N = \{1, 2, 3, 4\}$ and $a = \{10, 9, 0.5, 0.1\}$. From Steps 1–4, $S_1^e = \{1, 2\}$ and $S^m = \{3, 4\}$ are formed, where $u_1^e(S_1^e) = u_2^e(S_1^e) = 180.5$, $u_3^m(S^m) = 0.3$ and $u_4^m(S^m) = 0.06$. If the grand meritocratic coalition is formed, individual 1's payoff increases to 196, whereas individual 2's payoff decreases to 176.4. If only individual 1 forms a meritocratic coalition with individuals 3 and 4, then her/his payoff is 106, which is lower than the egalitarian payoff in $S_1^e = \{1, 2\}$, that is, 180.5. It is not profitable for individual 2 to join the meritocratic coalition $S^m$, either. Therefore, the structure is stable.

Though it does not affect our main results, it may be of interest to examine whether a stable structure exists in which individuals with higher abilities belong to a meritocratic coalition and individuals with lower abilities belong to an egalitarian coalition. The following example shows, in a bit of an abstract way, that this case can possibly occur.

**Example 4.** Consider $N = \{1, 2, 3, 4\}$ and $a = \{a_1, a_2, a_3, a_4\}$. Suppose that, as a result of the procedure, individual 1 belongs to a singleton and the other three individuals form an egalitarian coalition $S$. That is,

$$f(a_2), \frac{f(a_2 + a_3)}{2} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3} < \frac{f\left(\sum_{i\in N} a_i\right)}{4} \le f(a_1), \tag{3}$$

where the second inequality follows from Lemma 1.

For this structure to be stable, the meritocratic payoffs for individuals 2 to 4, if they form a meritocratic coalition with individual 1, must be weakly lower than their egalitarian payoff in $S$:

$$a_2 \times \frac{f(a_1 + a_2)}{a_1 + a_2} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3} = \bar{a}_S \times \frac{f\left(\sum_{i\in S} a_i\right)}{\sum_{i\in S} a_i} \Leftrightarrow \frac{f(a_1 + a_2)}{a_1 + a_2} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3a_2} \tag{4}$$

$$a_3 \times \frac{f\left(\sum_{i=1}^{3} a_i\right)}{\sum_{i=1}^{3} a_i} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3} = \bar{a}_S \times \frac{f\left(\sum_{i\in S} a_i\right)}{\sum_{i\in S} a_i} \Leftrightarrow \frac{f\left(\sum_{i=1}^{3} a_i\right)}{\sum_{i=1}^{3} a_i} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3a_3} \tag{5}$$

$$a_4 \times \frac{f\left(\sum_{i\in N} a_i\right)}{\sum_{i\in N} a_i} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3} \Leftrightarrow \frac{f\left(\sum_{i\in N} a_i\right)}{\sum_{i\in N} a_i} \le \frac{f\left(\sum_{i\in S} a_i\right)}{3a_4} \tag{6}$$

From Equation (4), $a_2 > a_3 > a_4$ and $a_2 > \bar{a}_S$, we have $a_1 < a_3 + a_4 < 2a_3$, or $a_4 > a_1 - a_3$ and $a_1/2 < a_3 < a_2$. From Equation (5) and $\sum_{i=1}^{3} a_i > \sum_{i\in S} a_i$, we have $a_3 < \bar{a}_S$, or equivalently, $a_2 - a_3 > a_3 - a_4$ must hold.

If all the conditions (3)–(6) are satisfied, the structure is stable.

## 3.4 Ability and payoff

Lastly, we note a general property of the relation between ability and payoff that holds for each case of decreasing, constant, or increasing marginal productivities. As in Farrell and Scotchmer (1988), wherein total surplus is always divided in an egalitarian way, the payoff of an individual in a stable structure is (at least weakly) greater than the payoffs of those with lower abilities in the current setting as well.

**Proposition 7.** *In a stable structure, each individual's payoff is at least weakly higher than the payoffs of individuals with lower abilities.*

*Proof.* From Propositions 1 and 2, the assertion is obviously true if $f'' = 0$ or $f'' < 0$. Suppose $f'' > 0$ and consider individuals $j$ and $k$ with $j < k$. If these individuals belong to the same coalition, then the claim follows directly. So let $j \in S$ and $k \in T$, where $S \neq T$.

(a) If $u_j^e(S) \leq u_k^e(T)$ or $u_j^m(S) \leq u_k^e(T)$, then the egalitarian coalition $(T\backslash\{k\})\cup\{j\}$ blocks the structure.

(b) If $u_j^e(S) \leq u_k^m(T)$, then the meritocratic coalition $T \cup \{j\}$ blocks the structure.

(c) If $u_j^m(S) \leq u_k^m(T)$, then the meritocratic coalition $S \cup T$ blocks the structure.

Therefore, individual $j$'s payoff cannot be lower than individual $k$'s payoff in any stable structure. □

The above proof demonstrates that if individuals $j$ and $k$ ($j < k$) belong to different coalitions in a stable structure, then the payoff of individual $j$ is *strictly* higher than that of individual $k$.

# 4 Discussion and conclusion

This paper analyzes a simple coalition formation game under different marginal productivities when total surplus can be distributed in either an egalitarian or a meritocratic manner. As can be easily expected, no individual forms a coalition with any other individuals if marginal productivity is decreasing, whereas if marginal productivity is constant, there is no egalitarian coalition in a stable structure and individuals are indifferent regarding the meritocratic coalition to which they belong, including singletons.

If marginal productivity is increasing, the grand meritocratic coalition is always stable. In addition, if the average total surplus is maximized by the grand coalition, then the grand egalitarian coalition is also stable. More generally, we propose a procedure to help identify stable structures, in which egalitarian and meritocratic coalitions may coexist. In stable structures, multiple egalitarian coalitions can coexist, all of which must always be consecutive, but there is, at most, one meritocratic coalition, which may or may not be consecutive.

This paper does not delve into the relationship between the distribution of abilities, marginal productivity, and stable structures. As the following examples demonstrate, consideration of more concrete production functions and ability distribution could present another research direction that will lead to a more comprehensive understanding of the effects of marginal productivity and ability distributions on stable structures.

**Example 5.** Suppose that $f(x) = x^2$ and $N = \{1, 2, 3, 4\}$. If the distribution of ability is more equal, such as $a = \{10, 9, 8, 7\}$, then the grand egalitarian coalition is formed according to the procedure. (From Proposition 3, the grand meritocratic coalition is also stable.)

In contrast, if abilities are more unequally distributed, such as $a' = \{10, 4, 1, 0.4\}$, or the highest-ability individual has a much higher ability than the rest of the population, as in $a'' = \{10, 3, 2, 1\}$, then the grand meritocratic coalition is formed.

**Example 6.** Suppose $N = \{1, 2, 3\}$ and $a = \{15, 6, 4\}$. If $f(x) = x^2$, then the grand meritocratic coalition is the only stable structure. However, if $f(x) = x^3$, the grand egalitarian coalition is also stable. More generally, it is expected that as marginal productivity increases, egalitarian coalitions are more likely to be contained in a stable structure.

Finally, in our analysis, we assume that egalitarian and meritocratic divisions are the social norms of surplus distribution in society. The investigation of what kinds of distribution norms arise with different levels of marginal productivity and different ability distributions is an important issue that deserves further research.[7]

# References


Suryapratim Banerjee, Hideo Konishi, and Tayfun Snmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18:135–153, 2001.

Salvador Barberà, Carmen Beviá, and Clara Ponsatí. Meritocracy, egalitarianism and the stability of majoritarian organizations. *Games and Economic Behavior*, 91:237–257, 2015.

Anna Bogomolnaia and Matthew O. Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

Joseph Farrell and Suzanne Scotchmer. Partnerships. *Quarterly Journal of Economics*, 103 (2):279–297, 1988.

Joseph Greenberg and Shlomo Weber. Strong tiebout equilibrium under restricted preferences domain. *Journal of Economic Theory*, 38(1):101–117, 1986.

P. Jean-Jacques Herings, Riccardo D. Saulle, and Christian Seel. The last will be first, and the first last: Segregation in societies with relative pay-off concerns. *The Economic Journal*, 131(637):2119–2143, 2021.

Massimo Morelli and In-Uck Park. Internal hierarchy and stable coalition structures. *Games and Economic Behavior*, 96:90–96, 2016.


[7] Nishimura et al. (2017) study the joint evolution of coalition formation and coalition-value distribution with a three-person ultimatum game.

Hervé Moulin. Equal or proportional division of a surplus, and other methods. *International Journal of Game Theory*, 16(3):161–186, 1987.

Hervé Moulin. *Axioms of Cooperative Decision Making*. Cambridge University Press, Cambridge, 1988.

Hervé Moulin. *Fair Division and Collective Welfare*. MIT Press, Cambridge, 2003.

Takeshi Nishimura, Akira Okada, and Yasuhiro Shirata. Evolution of fairness and coalition formation in three-person ultimatum games. *Journal of Theoretical Biology*, 420:53–67, 2017.

Debraj Ray. *A Game-Theoretic Perspective on Coalition Formation*. Oxford University Press, Oxford, 2007.

Debraj Ray and Rajiv Vohra. Coalition formation. In H. Peyton Young and Shmuel Zamir, editors, *Handbook of Game Theory*, volume 4, chapter 5, pages 239–326. North-Holland, Amsterdam, 2014.

Ronny Razin and Michele Piccione. Coalition formation under power relations. *Theoretical Economics*, 4(1):1–15, 2009.

John E. Roemer. *Theories of Distributive Justice*. Harvard University Press, Cambridge, 1996.

John E. Roemer and Joaquim Silvestre. The proportional solution for economies with both private and public ownership. *Journal of Economic Theory*, 59(2):426–444, 1993.

Alison Watts. Formation of segregated and integrated groups. *International Journal of Game Theory*, 35:505–519, 2007.