

The Last will be First, and the First Last:  
Segregation in Societies with Relative Payoff Concerns  
Short title: The Last will be First, the First Last

P. Jean-Jacques Herings \*    Riccardo D. Saulle    Christian Seel

June 17, 2020

**Abstract**

This paper studies coalition formation among individuals who differ in productivity. We consider egalitarian societies in which coalitions split their surplus equally and individualistic societies in which coalitions split their surplus according to productivity. Preferences of coalition members depend on their material payoffs, but are also influenced by relative payoff concerns.

The stable partitions in both egalitarian and individualistic societies are segregated, i.e., individuals with adjacent productivities form coalitions. If some individuals are not part of a productive coalition, then these are the least productive ones for egalitarian societies and the most productive ones for individualistic societies.

*KEYWORDS: Group Formation, Segregation, Relative Payoff, Egalitarianism, Meritocracy, Social Environment.*

*JEL classification: C70, C71, D62.*

---

\*Department of Economics, Maastricht University, P.O. Box 616, 6200 MD Maastricht, The Netherlands.  
E-Mail: P.Herings@maastrichtuniversity.nl

# 1 Introduction

Historically, the exercise of power has been pervasive to every society albeit being present in different forms. Those forms of coercion can be naturally traced back to the idea of the “state of nature”, a hypothetical primitive scenario in which neither property rights nor socio-political institutions of enforcement exist. According to the political philosopher Thomas Hobbes (1651), this condition of “mere nature” induces a *bellum omnium contra omnes* in which the most advantaged individuals are the physically strongest ones.

In the state of nature, however, individuals share the wish of ending the conflicts in order to claim properties and to benefit from their goods. This common desire leads into a “social contract” which bans the coercive force in favour of the enforcement power of institutions which ensure social norms and property rights.<sup>1</sup> In broad economic terms, different societies can be seen as the outcome of different social contracts.

In this paper, we study a frequently observed phenomenon in all non-primitive societies, the formation of coalitions - groups of individuals with a common goal. Different examples include firms, groups that provide a local public good, political parties, and cartels. We analyze situations in which the social contract is such that membership in a coalition is voluntary, i.e., individuals cannot exert their power or abilities in order to subjugate others.<sup>2</sup>

We model two possible and alternative outcomes of a social contract, an egalitarian and an individualistic society. The individualistic society does not allow coercion, but there is no enforcement of resource redistribution within it, i.e., each member of a coalition receives a material payoff which is proportional to his or her own productivity. The egalitarian society also does not allow for coercive force, but guarantees an equal distribution of resources within a coalition through institutions or norms, i.e., each member of a coalition receives the same material payoff.

In addition to the material payoff, preferences are influenced by positional concerns. The phenomenon of positional concern, especially about income and consumption, was first elaborated by Duesenberry (1949) who proposed the “demonstration effect” to explain how a family’s consumption is influenced by the purchases of its neighbors. Based on this observation, he developed the “relative income hypothesis” which states that the level of consumption satisfaction of an individual depends on the individual’s relative rank in the society.

Following the same line of reasoning, Frank (1985) argues that positional concerns can explain many real world phenomena such as flatter intrafirm wage profiles in which low productivity workers and high productivity workers are paid more and less with respect to their

---

<sup>1</sup>On this point of view, see Muthoo (2004) and Hafer (2006) who study models explaining the rise of property rights from the state of nature.

<sup>2</sup>Models with different forms of coercion are due to Jordan (2006), Piccione and Rubinstein (2007), Acemoglu, Egorov, and Sonin (2008), Piccione and Razin (2009), and Jordan and Obadia (2015).

marginal contribution to compensate for the difference in relative ranking. More recently, Card, Mas, Moretti, and Saez (2012) have shown that high relative wages affect job satisfaction and Bracha, Gneezy, and Loewenstein (2015) have shown a positive correlation between the supply of labor and relative wages. Moreover, happiness is found to be significantly and negatively affected by a lower relative income (Ferrer-i-Carbonell, 2005; Luttmer, 2005; Clark, Frijters, and Shields, 2008). To sum up, it is widely accepted that people do not only care about their absolute wealth, but also about their relative ranking in society.

Motivated by these findings, we study a coalition formation model with positional concerns. We model the importance of the relative position in a coalition by the intragroup relative payoff. This measures the difference between the material payoff of an individual and the average material payoff of the coalition to which the individual belongs.

We employ two different solution concepts to analyze societies. First, the classic concept of the Core which consists of all partitions for which there is no coalition such that all members benefit from deviating. As a second solution concept, we consider the Myopic Stable Set as introduced in Demuynck, Herings, Saulle, and Seel (2019). In the setting of this paper, the Myopic Stable Set is the minimal set of partitions such that (i) from any partition in the set, no sequence of myopic coalitional improvements leaves the set and (ii) for any given initial partition, there is a sequence of myopic coalitional improvements which leads to an element in the set.

For both egalitarian and individualistic societies, we provide a full characterization of the Core and the Myopic Stable Set and show that both solution concepts coincide. In both societies, every productive coalition consists of members with adjacent productivities. In egalitarian societies, only the least productive individuals may not be part of a productive coalition, while in individualistic societies only the most productive individuals may not be part of a productive coalition. This finding might be counterintuitive at first sight, as one might expect that egalitarian societies benefit less productive and individualistic societies are to the advantage of more productive individuals.

The reasoning behind the predictions for egalitarian societies is as follows. First, note that in an egalitarian society there is no envy among individuals, i.e. the relative payoff of each individual in each coalition is the same. Thus, individuals prefer to stay with highly productive individuals in order to increase their material payoffs. This leads to top-down segregation. In contrast, the results for individualistic societies rely on the fact that in individualistic societies, the individuals with the lowest productivity are very attractive as coalition members, since they do not affect the material payoffs but yield a higher relative payoff to other members of the coalition. Since these incentives are mutual, groups of individuals with low productivity will form. Intuitively, a similar process repeats until only individuals with high productivity might remain in unproductive coalitions. This process leads to bottom-up segregation.

The analysis of the egalitarian case would remain the same in the absence of relative payoff concerns. Since the specification of payoffs in the egalitarian model eliminates any positional concerns, there would still be top-down segregation. This finding is shared with Farrell and Scotchmer (1988) who analyze properties of the Core of an egalitarian society without relative payoff concerns under slightly stronger assumptions. The existence of relative payoff concerns is crucial for our findings in the individualistic case. Otherwise, individuals would not care about the productivity levels of their coalition partners and would only be concerned by joining a coalition that has sufficiently many members. As a result, there would be no segregation in individualistic societies when relative payoff concerns are absent.

We provide a comparative welfare analysis between the two types of societies. When all individuals differ in productivity and there are sufficient complementarities in production, stable partitions in egalitarian societies are always at least as materially efficient as stable partitions in individualistic societies. In an egalitarian society, the most productive individuals are always part of a productive coalition, while this might not be the case for individualistic societies.

## 1.1 Related Literature

The paper is part of the vast literature on the theory of coalition formation (e.g. Ray, 2007; Ray and Vohra, 2014). Our model belongs to the class of hedonic games (e.g., Banerjee, Konishi, and Sönmez, 2001; Bogomolnaia and Jackson, 2002), which are non-transferable utility games in which agents have preferences over the coalitions to which they belong. In particular, there are no externalities caused by other coalitions; for an overview on that topic, see Kóczy (2018).

Our solution concepts, the Core and the Myopic Stable Set are myopic in the sense that agents do not anticipate that deviations by coalitions may trigger further deviations by other coalitions. As is shown in Demuynck, Herings, Saulle, and Seel (2019), a Myopic Stable Set exists under weak assumptions on the primitives. Another well-known myopic solution concept is the von Neumann-Morgenstern stable set as introduced in von Neumann and Morgenstern (1944). One of the drawbacks of this latter concept is that it may not exist. If a von Neumann-Morgenstern stable set exists, it is guaranteed to have a non-empty intersection with the Myopic Stable Set by Theorem 3.15 of Demuynck, Herings, Saulle, and Seel (2019).

There is also an extensive literature that focuses on farsighted solution concepts, where agents consider the entire chain of coalitional deviations that could occur, and only care about the utility provided at the very end of the deviation process. For some notable examples of the farsighted approach, the reader is referred to Chwe (1994), Herings, Mauleon, and Vannetelbosch (2009), Page and Wooders (2009), and Ray and Vohra (2015).

Our work is closely related to an increasing branch of literature devoted to the role of social norms in group formation (e.g. Watts, 2007; Piccione and Rabin, 2009; Barberà, Beviá, and Ponsatí, 2015; Morelli and Park, 2016; Beviá, Córchon, and Romero-Medina, 2017). Watts (2007) studies a model with a fixed number of groups and agents who either all want to be the highest status agent in their group or who all want to join the highest status group that they can join. Piccione and Rabin (2009) consider a more abstract partition function form game in which the relative strength of coalitions is represented by an exogenous “power relation”. Agents care only about their social ranking, which is determined by their own individual power and by the power of the coalition to which they belong.

As in Barberà, Beviá, and Ponsatí (2015) and Morelli and Park (2016), we consider a model of coalition formation in which agents are vertically differentiated in productivity and societies can use different sharing rules to allocate the surplus in a society to its members. Morelli and Park (2016) study how the degree of heterogeneity in the remuneration of the individuals affects the number of rival coalitions and they find a positive relation. Barberà, Beviá, and Ponsatí (2015) consider a model where members of society vote between egalitarianism and meritocracy as the guiding principle to distribute the joint benefits of cooperation. An important difference between both works and ours is that we incorporate positional concerns into our analysis. Given our emphasis on relative payoff concerns, our contribution can be also broadly related to the literature on social and interdependent preferences (e.g. Dufwenberg, Heidhues, Kirchsteiger, Riedel, and Sobel, 2011; Maccheroni, Marinacci, and Rustichini, 2012).

Notable contributions to the study of segregation and stability in group formation are Milchtaich and Winter (2002), Peşki and Szentes (2013), Goyal, Hernández, Martínez-Cánovas, Moisan, Muñoz-Herrera, and Sánchez (2018), and Kets and Sandroni (2019). Milchtaich and Winter (2002) present a model of group formation based on the assumption that individuals prefer to associate with people similar to them. For the case where individuals’ characteristics are one-dimensional, they give sufficient conditions for stable partitions to be segregating. Preferences for similar agents are implicit in Goyal, Hernández, Martínez-Cánovas, Moisan, Muñoz-Herrera, and Sánchez (2018), who show both theoretically and experimentally how segregation can arise in coordination games when agents, belonging to one of two types, can choose with whom to interact. Kets and Sandroni (2019) endogenize agents’ preferences for interacting with their own group by modeling the process by which agents take others’ perspective. In our approach, segregation arises without the assumption that individuals have homophilous preferences. Peşki and Szentes (2013) show how segregation can even arise when types are payoff-irrelevant in an equilibrium of a dynamic economy where agents condition their decisions on the payoff-irrelevant characteristics of their potential partners.

## 2 Model

In this section, we introduce a model of a society where individuals differ in their productivity. Individuals have the possibility to form coalitions, think for example about the formation of local communities or firms. The production technology of a coalition includes complementarities, i.e., it exhibits increasing returns to scale up to a certain threshold and constant returns to scale above this threshold. The surplus created by a coalition is increasing in the productivity of the coalition members.

Societies differ in how individuals split the total surplus of a coalition among its members. In an egalitarian society, the surplus is divided equally, while the surplus is divided in proportion to productivity in an individualistic society. Individuals rank coalitions based on their own material payoffs. In case two coalitions provide the same material payoff to an individual, then the coalition where the individual has a higher payoff relative to the other members of the coalition is preferred.<sup>3</sup> This models the empirically documented increased happiness associated with a higher relative income.

We now formally describe the building blocks of a *society*  $(N, \lambda, \tau)$ . The finite set  $N = \{1, \dots, n\}$  consists of all individuals in the society. A coalition  $S$  is a non-empty subset of  $N$  and the collection of all possible non-empty coalitions is denoted by  $\mathcal{N} = 2^N \setminus \{\emptyset\}$ . The collection of all partitions of a coalition  $S \in \mathcal{N}$  is denoted by  $\mathbb{P}(S)$ . The set  $\mathbb{P}(N)$  of all partitions of  $N$  is also denoted by  $\mathbb{P}$ . We write  $|\mathcal{P}|$  for the number of coalitions in a partition  $\mathcal{P} \in \mathbb{P}$ .

Let  $\lambda \in \mathbb{R}_{++}^N$  be the *productivity vector*, where  $\lambda_i$  represents the productivity level of individual  $i$ . Without loss of generality, the individuals in  $N$  are indexed in decreasing order of productivity, i.e.,  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ . Let some coalition  $S \in \mathcal{N}$  be given. An individual  $i \in S$  such that  $\lambda_i \geq \lambda_j$  ( $\lambda_i \leq \lambda_j$ ) for all  $j \in S$  is called a strong (weak) individual in  $S$ . For  $S = N$ , we drop the reference to the coalition and simply refer to an individual as a strong or a weak individual. We define the top individual in  $S$  as  $t(S) = \min_{i \in S} i$  and the bottom individual in  $S$  as  $b(S) = \max_{i \in S} i$ , so as the lowest and highest numbered individual in coalition  $S$ , respectively. Notice that  $t(S)$  is a strong and  $b(S)$  is a weak individual in  $S$ .

The *surplus*  $u(S)$  of coalition  $S \in \mathcal{N}$  is given by  $\tau(|S|) \sum_{i \in S} \lambda_i$ , where  $\tau$  is a non-negative function on  $\{1, \dots, n\}$  with  $\tau(n) = 1$ . The function  $\tau$  is weakly increasing, which captures complementarities in production. We define the *productivity threshold*  $\rho \in \{1, \dots, n\}$  as the smallest number  $q \in \{1, \dots, n\}$  such that  $\tau(q) > 0$  and we define the *maximum productivity threshold*  $\bar{\rho} \in \{1, \dots, n\}$  as the smallest number  $q \in \{1, \dots, n\}$  such that  $\tau(q) = 1$ . Notice that  $\tau(q) = 1$  for every  $q \geq \bar{\rho}$ . The *average surplus* in coalition  $S$  is denoted by  $\bar{u}(S) = u(S)/|S|$ .

Our production function extends both the specification used in Farrell and Scotchmer

<sup>3</sup>We relax the assumption of a lexicographic preference structure in Section 6.

(1988) and in Barberà, Beviá, and Ponsatí (2015). Farrell and Scotchmer (1988) look at the same functional form but require  $\tau$  to be strictly increasing, which implies  $\bar{\rho} = n$ , and require  $\tau(q)/q$  to be strictly decreasing for  $q$  sufficiently large. The functional form in Barberà, Beviá, and Ponsatí (2015) follows if we take  $\tau(q) = 0$  for every  $q < \rho$  and we set  $\rho = \bar{\rho}$ . The commonly studied stag hunt game, where a coalition exceeding a particular size is needed to hunt stag, provides an illustration for the latter specification.

Our next step is to specify how a coalition  $S \in \mathcal{N}$  distributes its surplus over the coalition members. We denote the *material payoff* of individual  $i \in S$  by  $m_i(S) \in \mathbb{R}$ . We consider two natural ways to divide the surplus. A division is *egalitarian* if each member  $i$  of a coalition  $S$  receives the same payoff. A division is *individualistic* if each individual  $i \in S$  is paid proportional to the own productivity. These restrictions could be legal requirements, but they can also be thought of as social norms within a society. If the material payoff of individual  $i \in S$  is determined according to the egalitarian division, then it is denoted by  $m_i^e(S)$ , whereas we use  $m_i^{\text{ind}}(S)$  for the individualistic division.

**Definition 2.1** (Egalitarian Society). A society  $(N, \lambda, \tau)$  is *egalitarian* if surplus division within coalitions is egalitarian, i.e., for all  $S \in \mathcal{N}$  and for all  $i \in S$ ,

$$m_i^e(S) = \bar{u}(S) = \tau(|S|) \frac{\sum_{i \in S} \lambda_i}{|S|}.$$

**Definition 2.2** (Individualistic Society). A society  $(N, \lambda, \tau)$  is *individualistic* if surplus division within coalitions is individualistic, i.e., for all  $S \in \mathcal{N}$  and for all  $i \in S$ ,

$$m_i^{\text{ind}}(S) = \tau(|S|) \lambda_i.$$

We now incorporate relative payoff concerns for our given society  $(N, \lambda, \tau)$ . The *relative payoff* of individual  $i \in N$  in coalition  $S \in \mathcal{N}$  equals  $r_i(S) = m_i(S) - \bar{u}(S)$ . The relative payoff measures the material payoff of individual  $i$  relative to the average payoff in the individual's coalition. The relative payoff of individual  $i$  in coalition  $S$  in an individualistic society is denoted by  $r_i^{\text{ind}}(S)$ .<sup>4</sup> In an egalitarian society, the relative payoff of individual  $i$  in coalition  $S$  is equal to  $r_i^e(S) = 0$ .

We extend the definitions of material payoffs and relative payoffs from coalitions to partitions. For every  $\mathcal{P} \in \mathbb{P}$  and  $i \in N$ , we denote by  $m_i(\mathcal{P})$  the material payoff of individual  $i$  in partition  $\mathcal{P}$ , i.e.,  $m_i(\mathcal{P}) = m_i(S_i(\mathcal{P}))$ , where  $S_i(\mathcal{P})$  is the coalition to which individual  $i$  belongs in partition  $\mathcal{P}$ . Similarly, we write  $r_i(\mathcal{P}) = r_i(S_i(\mathcal{P}))$  for the relative payoff of individual  $i$  in partition  $\mathcal{P}$ .

<sup>4</sup>Our results for individualistic societies would extend to alternative definitions of relative payoff as long as an individual's relative payoff within a coalition increases in case a strong individual in the coalition leaves as well as in case a more productive individual is replaced by a less productive one.

We are now ready to define the preferences of individual  $i \in N$  over the partition space. For every  $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ , we define *relative payoff preferences*  $\succeq_i$  over  $\mathbb{P}$  by setting  $\mathcal{P} \succeq_i \mathcal{Q}$  if and only if either  $m_i(\mathcal{P}) > m_i(\mathcal{Q})$  or  $[m_i(\mathcal{P}) = m_i(\mathcal{Q}) \text{ and } r_i(\mathcal{P}) \geq r_i(\mathcal{Q})]$ . Thus, individual  $i$  prefers the partition with the higher material payoff and if two partitions deliver the same material payoff, the one with the higher relative payoff. We denote the asymmetric part of  $\succeq_i$  by  $>_i$ .

The final part of this section considers our modeling of the coalition formation process. Let us first illustrate our model informally in the context of firms. Workers can deviate by either joining another firm or by grouping together to found a new firm. In both cases, such a deviation results in a new coalition. If another firm has no workers involved in such a deviation, it remains intact and operates as before. If a firm is left by some workers, the firm continues to operate with the remaining workers. This uniquely determines the assignment of workers to firms which results after a deviation by some workers.

Formally, the final building block of our social environment is an *effectivity correspondence*  $E$ . For each pair of partitions  $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ , the effectivity correspondence  $E$  associates a collection of coalitions  $E(\mathcal{P}, \mathcal{Q}) \subseteq \mathcal{N}$  which can move from  $\mathcal{P}$  to  $\mathcal{Q}$ . As is common in this literature, only a single coalition moves at the same time.<sup>5</sup> If  $E(\mathcal{P}, \mathcal{Q}) = \emptyset$ , then no coalition can move from  $\mathcal{P}$  to  $\mathcal{Q}$ . If  $T \in E(\mathcal{P}, \mathcal{Q})$ , then coalition  $T$  can move from partition  $\mathcal{P}$  to partition  $\mathcal{Q}$ .

We distinguish three groups that are affected by a move of coalition  $T$  from  $\mathcal{P}$  to  $\mathcal{Q}$ : the moving individuals  $T$  themselves, the residual individuals, and the outsiders. The moving individuals  $T$  create a new coalition and leave their former coalition members, the residual individuals, behind. Formally, for every  $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$  and  $T \in E(\mathcal{P}, \mathcal{Q})$ , the residual individuals are given by

$$R(\mathcal{P}, T) = \{i \in N \setminus T \mid S_i(\mathcal{P}) \cap T \neq \emptyset\}.$$

The outsider coalitions  $\mathcal{U}$  are coalitions that are not influenced by the activity of the moving individuals. Formally, we have

$$\mathcal{U}(\mathcal{P}, T) = \{S \in \mathcal{P} \mid S \cap T = \emptyset\}.$$

The members of the coalitions in  $\mathcal{U}(\mathcal{P}, T)$  are called the outsiders and are collected in the set  $U(\mathcal{P}, T) = \bigcup_{S \in \mathcal{U}(\mathcal{P}, T)} S$ .

The related literature entertains different assumptions regarding the residual individuals. We consider the  $\delta$ -*model* (see Hart and Kurz, 1983) which prescribes that the residual

<sup>5</sup>This assumption does not affect the Core as each member of each moving coalition would have to be better off. In Section 5, where the Myopic Stable Set is studied, multiple coalitions are allowed to move sequentially.



individuals stay together.<sup>6</sup>

We also assume that the deviating coalition does not have the power to influence the outsiders. The effectivity correspondence is therefore defined as

$$E(\mathcal{P}, \mathcal{Q}) = \{T \in \mathcal{N} \mid T \in \mathcal{Q}, \mathcal{U}(\mathcal{P}, T) \subseteq \mathcal{Q}, \forall i \in R(\mathcal{P}, T), S_i(\mathcal{P}) \setminus T \in \mathcal{Q}\}, \quad \mathcal{P}, \mathcal{Q} \in \mathbb{P}.$$

It follows from the definition that, for every  $\mathcal{P} \in \mathbb{P}$ ,  $E(\mathcal{P}, \mathcal{P}) = \mathcal{P}$ , that is, any coalition in  $\mathcal{P}$  can choose to stay at  $\mathcal{P}$ . We write  $\mathcal{Q}(\mathcal{P}, T) \in \mathbb{P}$  to denote the partition that results when the current partition is  $\mathcal{P}$  and the moving coalition is  $T$ .

This completes the description of the effectivity correspondence and thereby of our social environment  $\Gamma = (N, \mathbb{P}, E, \{\succeq_i\}_{i \in N})$  induced by a society  $(N, \lambda, \tau)$ .

### 3 Stability and Segregation

In this section, we introduce the solution concept of the Core and our notion of segregation. To define the Core, we first introduce the concept of strict dominance which requires that each member of a deviating coalition is strictly better off.

**Definition 3.1** (Strict Dominance). Let  $\Gamma = (N, \mathbb{P}, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. A partition  $\mathcal{Q} \in \mathbb{P}$  *strictly dominates*  $\mathcal{P} \in \mathbb{P}$  by  $S \in \mathcal{N}$  if  $\mathcal{Q} = \mathcal{Q}(\mathcal{P}, S)$  and  $\mathcal{Q} \succ_i \mathcal{P}$  for all  $i \in S$ .

Let a coalition  $S \in \mathcal{N}$  be given. Let  $f_S : \mathbb{P} \rightarrow \mathbb{P}$  be the coalitional dominance correspondence, where, for every  $\mathcal{P} \in \mathbb{P}$ ,  $f_S(\mathcal{P})$  denotes the set of partitions that strictly dominate  $\mathcal{P}$  by  $S$ . Since the partition that results from  $E$  when the current partition is  $\mathcal{P}$  and the moving coalition is  $S$  is uniquely determined and equal to  $\mathcal{Q}(\mathcal{P}, S)$ , we have that  $f_S(\mathcal{P})$  is equal to the singleton  $\{\mathcal{Q}(\mathcal{P}, S)\}$  if  $\mathcal{Q}(\mathcal{P}, S) \succ_i \mathcal{P}$  for all  $i \in S$  and  $f_S(\mathcal{P})$  is equal to the empty set otherwise. We write  $f(\mathcal{P}) = \bigcup_{S \in \mathcal{N}} f_S(\mathcal{P})$  to denote the subset of  $\mathbb{P}$  consisting of all partitions that strictly dominate  $\mathcal{P}$ . Such partitions are also referred to as myopic improvements upon  $\mathcal{P}$ . The Core equals the set of partitions which are not strictly dominated by another partition.

**Definition 3.2** (Core). Let  $\Gamma = (N, \mathbb{P}, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. The *Core* equals the set of partitions  $C = \{\mathcal{P} \in \mathbb{P} \mid f(\mathcal{P}) = \emptyset\}$ .

<sup>6</sup>An alternative specification is the  $\gamma$ -model, which prescribes that residual individuals fall apart into singleton coalitions. Kóczy and Lauwers (2007) use this specification when defining the Minimal Dominant Set. In our context of a society, the  $\delta$ -model seems more natural to us as one would not expect a firm to fall apart into singletons if some individuals leave. Nevertheless, use of the  $\gamma$ -model would lead to similar results in our set-up.

We now give a formal definition of segregation.

**Definition 3.3** (Segregated Partition). Let  $(N, \lambda, \tau)$  be a society. The partition  $\mathcal{P} \in \mathbb{P}$  is *segregated* if the following two properties hold:

- i. **Coalitional assortativity:** For every  $S, T$  such that  $|S| \geq \rho$  and  $S \neq T$ , we have  $\lambda_{b(S)} \geq \lambda_{t(T)}$  or  $\lambda_{b(T)} \geq \lambda_{t(S)}$ .
- ii. **Homophily:** If  $S \in \mathcal{P}$  satisfies  $|S| > \bar{\rho}$ , then, for all  $i, j \in S$ , we have  $\lambda_i = \lambda_j$ .

Coalitional assortativity requires that each productive coalition contains only individuals with adjacent productivity.<sup>7</sup> Thus, our notion of segregation can be seen as socio-economic segregation where groups of similarly educated and able individuals form. Homophily requires that if a coalition strictly exceeds the maximum productivity threshold, then all individuals in the coalition have the same productivity. We include homophily as part of segregation as it restricts the group size for groups which are not completely homogeneous.

The following example illustrates the concepts introduced in this section.

**Example 3.4.** Consider an individualistic society  $(N, \lambda, \tau)$  with  $N = \{1, \dots, 7\}$  and productivity levels given by  $\lambda_1 = \lambda_2 = \lambda_3 > \lambda_4 = \dots = \lambda_7$ . We assume that one needs two individuals to reach maximum productivity, whereas singletons are not productive at all, so  $\tau(1) = 0$  and  $\tau(2) = 1$ . The partition  $\mathcal{P} = \{\{1, 3\}, \{2\}, \{4, 5, 6, 7\}\}$  is an example of a segregated partition. Coalitional assortativity and homophily are easily verified to be satisfied as each coalition consists of equally productive members.

Let us now check if some partition strictly dominates  $\mathcal{P} = \{\{1, 3\}, \{2\}, \{4, 5, 6, 7\}\}$ . All individuals in  $\{4, 5, 6, 7\}$  obtain their highest possible material payoff of  $\lambda_4$  and their highest possible relative payoff of 0. Thus, none of these individuals can be strictly better off in a different coalition. Individuals in the coalition  $\{1, 3\}$  can also not be part of a profitable deviation, as they obtain the highest possible material payoff of  $\lambda_1$  and cannot increase their relative payoff as no individual  $i \geq 4$  can be involved in a profitable deviation. It follows that individual 2 cannot find another individual to form a deviating coalition. Thus, the partition  $\mathcal{P} = \{\{1, 3\}, \{2\}, \{4, 5, 6, 7\}\}$  is in the Core.

<sup>7</sup>The coalition  $S$  in the definition of coalitional assortativity is restricted to be productive. The coalitions that belong to some segregated partition may therefore not form a convex geometry as defined in Bilbao, Jiménez, and López (1998). Moreover, even productive coalitions need not be consecutive/connected as defined in Greenberg and Weber (1986) and Ambec and Ehlers (2008), since productive coalitions in a segregated partition do not necessarily consist of consecutively numbered individuals. Example 3.4 presents a segregated partition with a coalition that is not consecutive.

## 4 Results for Egalitarian and Individualistic Societies

In this section, we study the general properties of egalitarian and individualistic societies with relative payoff preferences. The following procedure turns out to be crucial in finding Core elements.

**Definition 4.1** (Choice of the Strongest (Weakest) Procedure). Choose a strong (weak) individual and one of the individual's most preferred coalitions. Add that coalition as an element of the partition. Repeat the same process with a strong (weak) individual out of those not assigned to a coalition in a previous step and choose one of the individual's most preferred coalitions among those consisting of all individuals which have not yet been assigned to a coalition. Repeat the process until all individuals are assigned to a coalition within the partition.

Let us provide some intuition as to why the Choice of the Strongest Procedure helps us in finding Core partitions in egalitarian societies. In egalitarian societies, in each coalition, all individuals get the same share of the surplus. Thus, members of a coalition with the highest possible average surplus cannot be better off in any alternative coalition. Iterating the argument, we obtain the following proposition.

**Proposition 4.2.** *Consider an egalitarian society. A partition is in the Core if and only if it is an outcome of the Choice of the Strongest Procedure.*

Proposition 4.2 is closely related to the main result in Farrell and Scotchmer (1988) who analyze properties of the Core of an egalitarian society under slightly stronger assumptions.

For individualistic societies and a given coalition size, each member prefers the other members to be as unproductive as possible: the own material payoff does not change, but the relative payoff increases when other coalition members are less productive. Intuitively, this gives the weak individuals the power to choose their preferred coalition.

**Proposition 4.3.** *Consider an individualistic society. A partition is in the Core if and only if it is an outcome of the Choice of the Weakest Procedure.*

We define Algorithms A.5 and A.7 in the Appendix. These algorithms generate all Core partitions in individualistic and egalitarian societies, respectively. In either case, the common characteristic is that individuals with similar productivity levels group together, where the most attractive group members are the strongest individuals for egalitarian societies and the weakest individuals for individualistic societies. We thus obtain our main segregation result.

**Theorem 4.4.** *Let  $\Gamma = (N, \mathbb{P}, E, \{\geq_i\}_{i \in N})$  be a social environment induced by an egalitarian or an individualistic society  $(N, \lambda, \tau)$ . The Core is non-empty. Each partition in the Core of  $\Gamma$  is a segregated partition.*

Intuitively, in an egalitarian (individualistic) society, following the Choice of the Strongest (Weakest) Procedure, a strong (weak) remaining individual chooses to group together with other remaining strong (weak) individuals whenever there are sufficiently many other individuals left to form a productive coalition. Thus, segregation and the Choice of the Strongest (Weakest) immediately yield an important corollary: if individuals are not part of a productive coalition, these have to be the least (most) productive ones.

**Corollary 4.5.**

- (i) Consider a partition  $\mathcal{P}$  in the Core of an egalitarian society. For every  $i \in N$  such that  $|S_i(\mathcal{P})| < \rho$  it holds that  $\lambda_i \leq \lambda_{n-\rho+1}$ .
- (ii) Consider a partition  $\mathcal{P}$  in the Core of an individualistic society. For every  $i \in N$  such that  $|S_i(\mathcal{P})| < \rho$  it holds that  $\lambda_i \geq \lambda_\rho$ .

In the remainder, we discuss special cases in which the effects leading to the previous corollary can be seen in the cleanest way. In particular, we will obtain a unique prediction for the Core: segregated coalitions of size equal to the maximum productivity threshold  $\bar{\rho}$  form and in egalitarian (individualistic) societies at most one coalition with the least (most) productive individuals does not reach the maximum productivity threshold.

To formalize the result, we first introduce some additional notation. For  $i, j \in N$  with  $i \leq j$ , following Greenberg and Weber (1986), the coalition  $\{i, \dots, j\}$  is called a *consecutive coalition* and is denoted by  $[i, j]$ . Let  $\ell$  and  $z$  be non-negative integers such that  $z < \bar{\rho}$  and the number of individuals in the society equals  $n = \ell\bar{\rho} + z$ . The number  $\ell$  is equal to the highest possible number of maximally productive coalitions.

For  $k \in \{1, \dots, \ell\}$ , we define  $T^k = [\bar{\rho}(k-1)+1, \bar{\rho}k]$  and call it the  $k^{th}$  top-down segregated coalition. The agents in  $T^{\ell+1} = [n-z+1, n]$  are excluded from participating in a maximally productive coalition. If  $T^{\ell+1} = \emptyset$  then we define the partition  $\mathcal{T} = \{T^k \mid k = 1, \dots, \ell\}$  and if  $T^{\ell+1} \neq \emptyset$ , then we define the partition  $\mathcal{T} = \{T^k \mid k = 1, \dots, \ell\} \cup T^{\ell+1}$ . We call  $\mathcal{T}$  the top-down segregated partition.

For  $k \in \{1, \dots, \ell\}$ , we define  $B^k = [n - \bar{\rho}k + 1, n - \bar{\rho}(k-1)]$  and call it the  $k^{th}$  bottom-up segregated coalition. The agents in  $B^{\ell+1} = [1, z]$  are excluded from participating in a maximally productive coalition. If  $B^{\ell+1} = \emptyset$  then we define the partition  $\mathcal{B} = \{B^k \mid k = 1, \dots, \ell\}$  and if  $B^{\ell+1} \neq \emptyset$ , then we define the partition  $\mathcal{B} = \{B^k \mid k = 1, \dots, \ell\} \cup B^{\ell+1}$ . We call  $\mathcal{B}$  the bottom-up segregated partition.

We now present conditions such that individuals in egalitarian (individualistic) societies form the top-down (bottom-up) segregated partition. One condition is that no two individuals have exactly the same productivity.

**Definition 4.6** (Complete Heterogeneity). A society  $(N, \lambda, \tau)$  is *completely heterogeneous* if all individuals differ in their productivity, i.e., for all  $i, j \in N$  with  $i \neq j$ , we have  $\lambda_i \neq \lambda_j$ .

We first present a characterization of the Core for individualistic societies.

**Proposition 4.7.** *Let  $\Gamma = (N, \mathbb{P}, E, \{\geq_i\}_{i \in N})$  be the social environment induced by the completely heterogeneous and individualistic society  $(N, \lambda, \tau)$ . Let  $\tau$  be strictly increasing up to  $\bar{\rho}$ . The Core is equal to  $\{\mathcal{B}\}$ .*

For egalitarian societies, we need that consecutive coalitions do not increase their average surplus by excluding the bottom individual in the coalition.

**Definition 4.8** (Cohesiveness of Adjacent Groups). A society satisfies *cohesiveness of adjacent groups* if for every consecutive coalition  $[i, j]$  such that  $2 \leq j - i + 1 \leq \bar{\rho}$  it holds that  $\bar{u}([i + 1, j]) < \bar{u}([i, j])$ .

Cohesiveness of adjacent groups requires that complementarities are large enough, i.e., consecutive coalitions of size less than or equal to the maximum productivity threshold  $\bar{\rho}$  prefer to stay together rather dropping the weakest individual from the coalition. This property implies that  $\tau$  is strictly increasing up to  $\bar{\rho}$ .

In individualistic societies, productive groups always try to reach the maximum productivity threshold as this increases the material payoff of each member. In contrast, the cohesiveness condition imposes a much stronger restriction on partition formation in egalitarian societies: here, adding another member to reach the productivity threshold might decrease the material payoff without the cohesiveness condition. If an extra member is added, then even if  $\tau$  increases, the average payoff might decrease.

Under these stronger conditions, we obtain a unique prediction for the Core.

**Proposition 4.9.** *Let  $\Gamma = (N, \mathbb{P}, E, \{\geq_i\}_{i \in N})$  be the social environment induced by the completely heterogeneous and egalitarian society  $(N, \lambda, \tau)$  which satisfies cohesiveness of adjacent groups. The Core is equal to  $\{\mathcal{T}\}$ .*

Under these conditions, we also obtain a clean comparison of Core partitions in both societies in terms of the resulting material efficiency.

**Definition 4.10** (Material Efficiency). Let  $(N, \lambda, \tau)$  be a society. The *material efficiency* of a partition  $\mathcal{P} \in \mathbb{P}$  is equal to  $\sum_{i \in N} m_i(\mathcal{P})$ .

The maximal material efficiency a society can reach is given by  $\sum_{i \in N} \lambda_i$ , which is for instance attained if the grand coalition forms. We now compare egalitarian and individualistic societies in terms of the resulting material efficiency.

**Theorem 4.11.** *Let  $(N, \lambda, \tau)$  be a completely heterogeneous society which satisfies cohesiveness of adjacent groups. The unique Core partition  $\mathcal{T}$  of the social environment induced by the egalitarian society is at least as materially efficient as the unique Core partition  $\mathcal{B}$  of the social environment induced by the individualistic society. They are equally materially efficient if and only if  $n$  is an integer multiple of  $\bar{\rho}$ .*

Only the strongest individuals are excluded from a maximally productive coalition in individualistic societies, while only the weakest individuals are excluded from a maximally production coalition in egalitarian societies. Thus, the associated loss in material efficiency is higher when stronger individuals do not reach their maximal productivity.

## 5 The Myopic Stable Set

The Core is a static solution concept. Once a Core partition is reached, no coalition has an incentive to deviate. But, starting from an arbitrary partition in the society, it is not clear that a partition in the Core will always be reached. In this section, we introduce the Myopic Stable Set. In the present setting with a finite partition space, the Myopic Stable Set coincides with the Core whenever for any partition outside the Core, there exists a sequence of myopic improvements which leads to a partition inside the Core. In such cases, it provides an additional dynamic foundation for the prediction of the Core.

To define the Myopic Stable Set, it is convenient to define, for every  $\mathcal{P} \in \mathbb{P}$ ,  $\tilde{f}(\mathcal{P}) = f(\mathcal{P}) \cup \{\mathcal{P}\}$ , so  $\tilde{f}(\mathcal{P})$  is obtained from  $f(\mathcal{P})$  by adding the partition  $\mathcal{P}$ . We define the two-fold composition of  $\tilde{f}$  by  $\tilde{f}^2(\mathcal{P}) = \{\mathcal{R} \in \mathbb{P} \mid \exists \mathcal{Q} \in \tilde{f}(\mathcal{P}) : \mathcal{R} \in \tilde{f}(\mathcal{Q})\}$ . For  $k \in \mathbb{N}$ , we define the  $k$ -fold iteration  $\tilde{f}^k(\mathcal{P})$  by induction as  $\tilde{f}^k(\mathcal{P}) = \{\mathcal{R} \in \mathbb{P} \mid \exists \mathcal{Q} \in \tilde{f}^{k-1}(\mathcal{P}) : \mathcal{R} \in \tilde{f}(\mathcal{Q})\}$ . We define the set of all partitions that can be reached from  $\mathcal{P}$  by a finite number of dominations by  $\tilde{f}^{\mathbb{N}}(\mathcal{P}) = \bigcup_{k \in \mathbb{N}} \tilde{f}^k(\mathcal{P})$ .

If  $k \leq \ell$ , then  $\tilde{f}^k(\mathcal{P}) \subseteq \tilde{f}^\ell(\mathcal{P})$ . Because the set  $\mathbb{P}$  is finite, it therefore holds that there is  $k' \in \mathbb{N}$  such that, for every  $k \geq k'$ ,  $\tilde{f}^k(\mathcal{P}) = \tilde{f}^{k'}(\mathcal{P})$ , so  $\tilde{f}^{\mathbb{N}}(\mathcal{P}) = \tilde{f}^{k'}(\mathcal{P})$ .

We are now ready to introduce the Myopic Stable Set, following the definition of Demuyneck, Herings, Saulle, and Seel (2019) for general social environments.<sup>8</sup>

**Definition 5.1** (Myopic Stable Set). Let  $\Gamma = (N, \mathbb{P}, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. The set  $M \subseteq \mathbb{P}$  is a *Myopic Stable Set* if it satisfies the following three conditions:

1. **Deterrence of external deviations:** For every partition  $\mathcal{P} \in M$ ,  $f(\mathcal{P}) \subseteq M$ .
2. **Iterated external stability:** For every  $\mathcal{Q} \in \mathbb{P} \setminus M$ , we have that  $\tilde{f}^{\mathbb{N}}(\mathcal{Q}) \cap M \neq \emptyset$ .
3. **Minimality:** There is no proper subset  $M'$  of  $M$  that satisfies Conditions 1 and 2.

<sup>8</sup>Demuyneck, Herings, Saulle, and Seel (2019) allow for social environments with an infinite state space. For the finite case, their closedness requirement is automatically satisfied and  $\tilde{f}^\infty(\mathcal{Q})$ , the closure of  $\tilde{f}^{\mathbb{N}}(\mathcal{Q})$ , can be replaced by  $\tilde{f}^{\mathbb{N}}(\mathcal{Q})$  in the definition of iterated external stability. Thus, for the finite case, the definition of Myopic Stable Set in Demuyneck, Herings, Saulle, and Seel (2019) is equivalent to Definition 5.1.

The first condition requires that no partition in the set is dominated by a partition outside the set. By the second condition, from any partition outside the set, there is a finite sequence of dominations which reaches the set. The final condition requires the set to be a minimal set which satisfies the first two conditions.

By Theorem 3.1 and Corollary 3.11 in Demuyne, Herings, Saulle, and Seel (2019) a Myopic Stable Set (MSS) exists and the Core is a subset of the MSS. Moreover, if the dominance correspondence  $\tilde{f}$  is lower hemi-continuous, then the MSS is unique by their Theorem 3.4. Lower hemi-continuity of  $\tilde{f}$  can be used to show that the intersection of two Myopic Stable Sets is a Myopic Stable Set itself, which establishes the uniqueness property. Notice that lower hemi-continuity is trivially satisfied for finite partition spaces. Thus, the MSS of the social environment induced by a society is always unique. The property of iterated external stability guarantees that the MSS is non-empty.

The MSS is conceptually related to the Minimal Dominant Set as described for transferable utility (TU) games in Kóczy and Lauwers (2007), but there are also some notable differences between these two notions. Contrary to Kóczy and Lauwers (2007) we require that when coalition members leave a coalition, the remaining individuals still form a coalition. We also impose a particular sharing rule for each coalition. Moreover, Kóczy and Lauwers (2007) consider weak improvements and require all payoff-equivalent states to be part of the Minimal Dominant Set.

We now characterize the MSS for our societies.

**Theorem 5.2.** *Let  $\Gamma = (N, \mathbb{P}, E, \{\geq_i\}_{i \in N})$  be a social environment induced by an egalitarian or individualistic society  $(N, \lambda, \tau)$ . The Core coincides with the Myopic Stable Set.*

The equivalence of the Core and the Myopic Stable Set gives an additional dynamic foundation for the prediction: from any partition, there is a path of myopic improvements to a segregated Core partition. We now illustrate such a path of myopic improvements for the society in Example 3.4.

**Example 5.3.** Consider an individualistic society with primitives as given in Example 3.4, i.e.,  $(N, \lambda, \tau)$  with  $N = \{1, \dots, 7\}$ ,  $\lambda_1 = \lambda_2 = \lambda_3 > \lambda_4 = \dots = \lambda_7$ ,  $\tau(1) = 0$ , and  $\tau(2) = 1$ . Let the initial partition be  $\mathcal{P}_0 = \{\{1\}, \{2, 3\}, \{4, 5, 6\}, \{7\}\}$ . Forming a coalition with individual 7 is a strict improvement for individual 2, since the relative payoff of individual 2 improves when individual 3 is replaced by individual 7. Since the material payoff improves for individual 7 by forming a productive coalition, forming a coalition with individual 2 is a strict improvement for individual 7 as well. The new partition is given by  $\mathcal{P}_1 = \{\{1\}, \{2, 7\}, \{3\}, \{4, 5, 6\}\}$ . In the next step, both individuals 1 and 3 increase their material payoffs by forming a coalition. The new partition is given by  $\mathcal{P}_2 = \{\{1, 3\}, \{2, 7\}, \{4, 5, 6\}\}$ , a partition in the Core.

For general social environments, the Myopic Stable Set can be strictly larger than the Core. For instance, this happens in all situations where the Core is empty, as the Myopic Stable Set always exists and is non-empty. One such example is provided in the Appendix for the case where preferences are determined by a convex combination of material and relative payoffs. The following example demonstrates that the Myopic Stable Set can also be a strict superset of the Core when the Core is non-empty.

**Example 5.4.** Consider a society with three agents and assume material payoffs to be given by  $m(\{\{1\}, \{2\}, \{3\}\}) = (0, 0, 0)$ ,  $m(\{\{1, 2\}, \{3\}\}) = (3, 1, 0)$ ,  $m(\{\{1\}, \{2, 3\}\}) = (0, 3, 1)$ ,  $m(\{\{1, 3\}, \{2\}\}) = (1, 0, 3)$ , and  $m(\{\{1, 2, 3\}\}) = (2, 2, 2)$ . The Core is given by  $\{\{1, 2, 3\}\}$ . Any partition where a two-agent coalition forms does not belong to the Core as it can be blocked by another two-agent coalition. Since it is impossible to reach the Core from such a partition by a path of myopic improvements, the Myopic Stable Set must be strictly larger than the Core. It is easily verified that the Myopic Stable Set is given by all partitions where a two-agent or a three-agent coalition forms. The material payoffs in this example cannot result in an egalitarian or individualistic society, so the example does not contradict Theorem 5.2.

## 6 Robustness

In this section, we relax our assumptions on the preferences and the production function.

### 6.1 Beyond Lexicographic Preferences

In the main text, we have assumed for analytical convenience that individuals have lexicographic preferences. Individuals care first about material payoffs, then about relative payoffs. We now consider how our results change in individualistic societies if we relax this assumption.<sup>9</sup> Suppose preferences can be represented by a convex combination of the material and relative payoffs, i.e., for every  $S \in \mathcal{N}$ , payoffs of  $i \in S$  are equal to  $p_i(S) = \mu m_i^{\text{ind}}(S) + (1 - \mu)r_i^{\text{ind}}(S)$  for some given  $\mu \in (0, 1)$ . This can be rewritten as

$$\begin{aligned} p_i(S) &= \mu m_i^{\text{ind}}(S) + (1 - \mu)r_i^{\text{ind}}(S) \\ &= \mu m_i^{\text{ind}}(S) + (1 - \mu)(m_i^{\text{ind}}(S) - \bar{u}(S)) \\ &= m_i^{\text{ind}}(S) - (1 - \mu)\bar{u}(S). \end{aligned}$$

First note that for any given parameters, there exists an  $\epsilon > 0$  such that the results in the main text extend to  $\mu \in (1 - \epsilon, 1)$ : for  $\mu$  sufficiently large, when the material payoffs differ, individual  $i$  always prefers the coalition with the higher material payoff. Only in

<sup>9</sup>In egalitarian societies, relative payoffs are always equal to zero and results remain the same.



case of equal material payoffs,  $i$  prefers the coalition with the higher relative payoffs which reconstructs the preferences from the main text.

Second, for any given parameters, there also exists an  $\epsilon > 0$ , such that for  $\mu \in (0, \epsilon)$  each individual prefers the coalition with the highest relative payoff. In particular, for a completely heterogeneous and individualistic society, this leads to only singleton coalitions, as no individual wants to be the least productive member of a coalition.

For intermediate values of  $\mu$ , the Core might be empty. In the Appendix we provide an example with an empty Core and we derive the Myopic Stable Set for that example.

## 6.2 A Different Production Function

In the main text, the function  $\tau$  depends on the size of each coalition. Instead, this function might depend on the sum of the individual productivities, i.e.,  $\tau(S) = \tau(\sum_{i \in S} \lambda_i)$  with a surplus  $u(S) = \tau(\sum_{i \in S} \lambda_i) \sum_{i \in S} \lambda_i$ . In this case, the results are similar, but less clean as the following example illustrates.

**Example 6.1.** Consider an individualistic society. Suppose  $\tau(S) = 1$  if  $\sum_{i \in S} \lambda_i \geq 10$  and  $\tau(S) = 0$  otherwise. Let  $n = 8$ ,  $\lambda_1 = 5.5$ ,  $\lambda_2 = 5.1$ ,  $\lambda_3 = 4.5$ ,  $\lambda_4 = 4.3$ ,  $\lambda_5 = 4.2$ ,  $\lambda_6 = 4.1$ ,  $\lambda_7 = 4$ ,  $\lambda_8 = 3$ ,  $\lambda_9 = 2$ , and  $\lambda_{10} = 1$ .

In this case, segregation starts bottom-up: The coalition  $\{7, 8, 9, 10\}$  yields the strictly highest material payoff and - conditional on being productive - the highest relative payoff for each member. Among the remaining individuals, the same is true for coalition  $\{4, 5, 6\}$ . Among the first three individuals, only coalitions involving individual 1 are productive. But the formation of coalition  $\{1, 2\}$  or  $\{1, 2, 3\}$  leads to a profitable deviation by coalition  $\{1, 3\}$ . In both cases, the relative payoff improves for individual 1 since an individual with above-average productivity is dropped while the coalition remains productive. The material payoff of individual 3 increases in the former case, whereas it remains constant while the relative payoff increases in the latter case. Thus, the unique Core element is  $\{\{1, 3\}, \{2\}, \{4, 5, 6\}, \{7, 8, 9, 10\}\}$ . This partition is not segregated as  $\lambda_1 > \lambda_2 > \lambda_3$  and yet the productive coalition  $\{1, 3\}$  is not a consecutive coalition.

Thus, as in the main text, a weak individual chooses a preferred coalition in each step. Yet, in contrast to the main text, a productive coalition is not necessarily segregated, as individuals who have above-average productivity and are not pivotal for a productivity increase might be dropped.

# Appendix

## A.1 Preliminaries

We establish a simple but useful result for the case in which some individuals have the same productivity level. Informally, if a partition in which these individuals are in different coalitions is in the Core, then all partitions in which these individuals swap their coalitions are also in the Core. To define this formally, for given  $\mathcal{P}, \mathcal{Q} \in \mathbb{P}$ , let  $h : N \rightarrow N$  be a *productivity preserving permutation*, i.e.,  $h$  is a one-to-one mapping such that  $h(i) = j$  implies  $\lambda_i = \lambda_j$ . Moreover, we define  $h(S) = \{j \in N \mid \exists i \in S \text{ such that } j = h(i)\}$  as the image of  $S$  under  $h$ .

**Definition A.1** (Equivalence). Let  $(N, \lambda, \tau)$  be a society. The partitions  $\mathcal{P}$  and  $\mathcal{Q}$  in  $\mathbb{P}$  are *equivalent* if there is a productivity preserving permutation  $h$  such that for all coalitions  $S \in \mathcal{P}$  it holds that  $h(S) \in \mathcal{Q}$ .

Since  $h$  can be the identity, any partition  $\mathcal{P}$  is equivalent to itself. We denote the equivalence class of a partition  $\mathcal{P}$  by  $[\mathcal{P}]$ . The next result shows if a partition is in the Core, then all equivalent partitions are also in the Core.

**Proposition A.2.** Let  $\Gamma = (N, \mathbb{P}, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by an egalitarian or an individualistic society. If  $\mathcal{P} \in C$ , then  $[\mathcal{P}] \subseteq C$ .

*Proof.* Assume  $\mathcal{P} \in C$  and  $\mathcal{Q} \in [\mathcal{P}]$ . Let  $h$  be a productivity preserving permutation such that for all coalitions  $S \in \mathcal{P}$  it holds that  $h(S) \in \mathcal{Q}$ . Suppose  $\mathcal{Q} \notin C$ . Then there exists  $\mathcal{R} \in \mathbb{P}$  and  $T \in \mathcal{N}$  such that  $\mathcal{R} \in f_T(\mathcal{Q})$ . We define  $h^{-1}(\mathcal{R}) = \{S \in \mathcal{N} \mid h(S) \in \mathcal{R}\}$ . It holds that  $h^{-1}(\mathcal{R}) \in f_{h^{-1}(T)}(\mathcal{P})$ , contradicting that  $\mathcal{P} \in C$ .  $\square$

We will show when proving Theorem 5.2 that the predictions of the Myopic Stable Set and the Core coincide. Thus, if the Myopic Stable Set contains a partition  $\mathcal{P}$ , it also contains all partitions in  $[\mathcal{P}]$ .

The following lemmas are helpful in the proof of Theorem 4.4. In a coalition where some individuals differ in productivity and where the size exceeds the maximum productivity threshold, the members are better off excluding the least (most) productive individual in an egalitarian (individualistic) society.

**Lemma A.3.** Let  $\Gamma = (N, \mathbb{P}, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by an egalitarian society  $(N, \lambda, \tau)$ . Let  $\mathcal{P} \in \mathbb{P}$  and  $S \in \mathcal{P}$  be such that  $|S| > \bar{\rho}$  and  $\lambda_i \neq \lambda_j$  for some  $i, j \in S$ . Then  $f_{S \setminus \{b(S)\}}(\mathcal{P}) \neq \emptyset$ .

*Proof.* Consider the coalition  $T = S \setminus \{b(S)\}$ . Since  $|T| \geq \bar{\rho}$ , we have  $u(S) = u(T) + \lambda_{b(S)}$ . Observe that  $\lambda_{b(S)} < \bar{u}(S)$  since  $\lambda_i \neq \lambda_j$  for some  $i, j \in S$ . Hence, we have  $\bar{u}(T) > \bar{u}(S)$ .  $\square$

**Lemma A.4.** *Let  $\Gamma = (N, \mathbb{P}, E, \{\geq_i\}_{i \in N})$  be a social environment induced by an individualistic society  $(N, \lambda, \tau)$ . Let  $\mathcal{P} \in \mathbb{P}$  and  $S \in \mathcal{P}$  be such that  $|S| > \bar{\rho}$  and  $\lambda_i \neq \lambda_j$  for some  $i, j \in S$ . Then  $f_{S \setminus \{t(S)\}}(\mathcal{P}) \neq \emptyset$ .*

*Proof.* Consider the coalition  $T = S \setminus \{t(S)\}$ . Since  $|T| \geq \bar{\rho}$ , we have  $m_i^{\text{ind}}(T) = m_i^{\text{ind}}(S)$  for all  $i \in T$ . Hence, to prove that  $f_{S \setminus \{t(S)\}} \neq \emptyset$ , it remains to show that  $r_i^{\text{ind}}(T) > r_i^{\text{ind}}(S)$  for all  $i \in T$ .

Observe that  $\lambda_{t(S)} > \bar{u}(S)$  since  $\lambda_i \neq \lambda_j$  for some  $i, j \in S$ . Hence, we have  $\bar{u}(T) < \bar{u}(S)$  and thus, for every  $i \in T$ ,  $r_i^{\text{ind}}(T) = m_i^{\text{ind}}(T) - \bar{u}(T) > m_i^{\text{ind}}(S) - \bar{u}(S) = r_i^{\text{ind}}(S)$ .  $\square$

## A.2 Proofs of the Results from the Main Text

*Proof of Proposition 4.2.* We first show that if a partition is an outcome of the Choice of the Strongest (henceforth CotS) Procedure, it is in the Core for an egalitarian society.

Any member of the coalition chosen by CotS in the first round obtains the highest possible material payoff. Thus, no such member can be involved in a myopic improvement. Iterate the argument for each of the following rounds of the CotS procedure to obtain the result.

We now show that any Core element for an egalitarian society can be obtained by the CotS procedure.

Take a partition in the Core. Note that the partition must contain at least one coalition with the highest possible material payoff per member as such a coalition would have a profitable deviation otherwise. As utility is increasing in the productivity of the group members, a strongest individual must be part of this group. Thus, this coalition can be picked up in the first round of the CotS procedure. Iterate the argument as long as the number of remaining individuals is greater than or equal to  $\rho$ . As soon as the number of remaining individuals is less than  $\rho$ , then all remaining coalitions generate a surplus equal to zero. Any strongest individual among the remaining individuals and his coalition can be picked up by the CotS procedure. Iterate the argument to obtain the result.  $\square$

*Proof of Proposition 4.3.* We first show that if a partition is an outcome of the Choice of the Weakest (henceforth CotW) Procedure, it is also in the Core for an individualistic society.

To do so, we first establish that each member of the coalition chosen by the weakest individual in the first round of the CotW procedure prefers that coalition over any other coalition. Since preferences are lexicographic, the weakest individual chooses a coalition which obtains the maximal possible value of  $\tau$  among all possible coalitions. Moreover, this coalition obtains the lowest possible value of  $\bar{u}$  among all coalitions with the maximal value of  $\tau$ . Therefore, all individuals in a preferred coalition of the weak individual weakly prefer that coalition over any other coalition. Thus, no such individual can be part of a deviating coalition. Iterating the argument with the remaining individuals leads to the result.

We now show that any Core element can be obtained by the CotW procedure.

Consider a partition in the Core. Towards a contradiction, suppose that no weakest individual is in one of the individual's most preferred coalitions. Consider a deviation by a coalition  $S$  that is the most preferred coalition of a weakest individual. Such a coalition has the highest possible value of  $\tau$  and, conditional on having the highest possible value of  $\tau$ , the lowest possible value of  $\bar{u}$ . This makes all weakest individuals in the coalition strictly better off. Moreover, all other members of that coalition receive a weakly higher material payoff as the weakest individual chooses a coalition with the maximal possible value of  $\tau$ . They are strictly better off if their material payoff is higher. Consider a non-weakest individual  $i \in S$  with the same material payoff as in the individual's original coalition, say  $T$ . If  $T$  contains some weakest individual, then  $\bar{u}(S) < \bar{u}(T)$  by the assumption that no weakest individual is in one of the individual's most preferred coalitions. If  $T$  contains no weakest individuals, then  $\bar{u}(S) < \bar{u}(T)$ , since the replacement of one member in  $T$  by a weakest individual leads to a lower value of  $\bar{u}$ , whereas  $S$  has the lowest such value conditional on having the highest possible value of  $\tau$ . In both cases, we have  $r_i^{\text{ind}}(S) < r_i^{\text{ind}}(T)$ . We conclude that all members of  $S$  are strictly better off. Thus, the partition cannot be in the Core. Consequently, in every Core element at least one weakest individual has to be in one of the individual's most preferred coalitions. Thus, this coalition can be picked up in the first step of the CotW procedure. Iterate the argument as long as the number of remaining individuals is greater than or equal to  $\rho$ . When the number of remaining individuals is less than  $\rho$ , then all remaining coalitions generate a surplus equal to zero. Any weakest individual among the remaining individuals and his coalition can be picked up by the CotW procedure. Iterate the argument to obtain the result.  $\square$

### An Algorithm to Find All Partitions in the Core

In what follows, we provide an algorithm to find all Core partitions of a given society. From the previous results, we know that a partition is in the Core of an individualistic society if and only if it can be picked up by the Choice of the Weakest Procedure. We now characterize all such partitions. We call  $[i - k, i]$  for  $k \in \{0, \dots, i - 1\}$  a *downward adjacent coalition of  $i$*  and  $[i, n - k]$  for  $k \in \{0, \dots, n - i\}$  an *upward adjacent coalition of  $i$* .

Our algorithm results in a weighted directed rooted tree  $(V, A, v^0)$ , where  $V$  is a set of decision nodes,  $A \subseteq V \times V$  is a set of ordered pairs of nodes called arcs, and  $v^0$  is a given initial decision node. At each decision node, it has to be decided how big a coalition to form. Each possible coalition size results in an arc from the current decision node to a successor decision node. To each arc  $a \in A$ , we associate an integer  $\mu_a \in \mathbb{N}$  called its weight. This weight  $\mu_a$  is equal to the size of the coalition that is formed at the current decision node. Given some  $v \in V$ , we denote by  $A^v \subseteq A$  the set of all arcs on the directed path from  $v^0$  to  $v$ . Since  $(V, A, v^0)$  is a directed rooted tree, the set  $A^v$  is uniquely determined, see Theorem 9A

of Wilson (1985). To each decision node  $v$ , we associate a number  $n(v)$ , the number of individuals that are not yet assigned to a coalition.

**Algorithm A.5.** Let  $(N, \lambda, \tau)$  be an individualistic society. Let  $v = v^0$  with  $v^0$  the initial decision node. The initial set of arcs  $A$  is equal to the empty set.

**Step 1** Define  $n(v) = n - \sum_{a \in A^v} \mu_a$ . If  $n(v) > 0$ , consider all downward adjacent coalitions of  $n(v)$ . Denote the number of downward adjacent coalitions which are most preferred by  $n(v)$  by  $k$ . Add  $k$  arcs starting at  $v$  and ending at distinct new nodes. The arcs all have different weights, each equal to the size of a most preferred downward adjacent coalition of  $n(v)$ . If  $n(v) = 0$ , then  $v$  is a terminal node and no new decision node is created.

**Step 2** For each new decision node  $v$  created in Step 1, go back to Step 1. Step 2 finishes when there are no new decision nodes anymore.

**Step 3** Consider the weighted rooted directed tree  $(V, A, v^0)$  resulting from Steps 1 and 2. For every terminal node  $v$ , let  $(a_1, \dots, a_\ell)$  be the arcs on the path from  $v^0$  to  $v$ , and generate the partition

$$\left\{ [1, n - \sum_{k=1}^{\ell-1} \mu_{a_k}], \dots, [n - \mu_{a_\ell} + 1, n] \right\}.$$

**Step 4** Collect all the partitions created in Step 3 and add all equivalent partitions.

We use Example 3.4 to illustrate Algorithm A.5. The weighted directed rooted tree generated in Steps 1 and 2 is illustrated in Figure 1. This directed rooted tree has five terminal nodes, resulting in five partitions in Step 3:

$$\begin{aligned} \mathcal{P}_1 &= \{\{1, 2, 3\}, \{4, 5\}, \{6, 7\}\}, \\ \mathcal{P}_2 &= \{\{1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}\}, \\ \mathcal{P}_3 &= \{\{1, 2\}, \{3, 4\}, \{5, 6, 7\}\}, \\ \mathcal{P}_4 &= \{\{1\}, \{2, 3\}, \{4, 5, 6, 7\}\}, \\ \mathcal{P}_5 &= \{\{1, 2, 3\}, \{4, 5, 6, 7\}\}. \end{aligned}$$

Step 4 adds equivalent partitions. By elementary combinatorics, this leads to two additional partitions for  $\mathcal{P}_1$ , eight additional partitions for  $\mathcal{P}_2$ , eleven additional partitions for  $\mathcal{P}_3$ , two additional partitions for  $\mathcal{P}_4$ , and no additional partitions for  $\mathcal{P}_5$ . None of these partitions is identical. Thus, we have a total of 28 segregated partitions which we list ordered by equivalence classes in the following table.

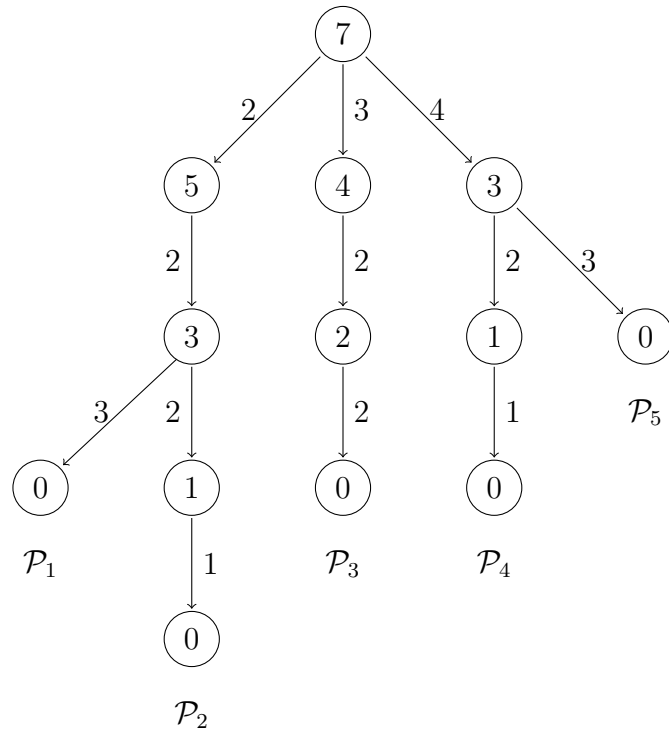


Figure 1: Result of Steps 1 and 2 of Algorithm A.5 for Example 3.4. The number  $n(v)$  is displayed in the circles, the numbers  $\mu_a$  are displayed on the arcs, and the generated partitions are indicated at the bottom.

$[\mathcal{P}_1]$	$\{\{1, 2, 3\}, \{4, 5\}, \{6, 7\}\}; \{\{1, 2, 3\}, \{4, 6\}, \{5, 7\}\}; \{\{1, 2, 3\}, \{4, 7\}, \{5, 6\}\}$
$[\mathcal{P}_2]$	$\{\{1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}\}; \{\{1\}, \{2, 3\}, \{4, 6\}, \{5, 7\}\}; \{\{1\}, \{2, 3\}, \{4, 7\}, \{5, 6\}\};$ $\{\{2\}, \{1, 3\}, \{4, 5\}, \{6, 7\}\}; \{\{2\}, \{1, 3\}, \{4, 6\}, \{5, 7\}\}; \{\{2\}, \{1, 3\}, \{4, 7\}, \{5, 6\}\};$ $\{\{3\}, \{1, 2\}, \{4, 5\}, \{6, 7\}\}; \{\{3\}, \{1, 2\}, \{4, 6\}, \{5, 7\}\}; \{\{3\}, \{1, 2\}, \{4, 7\}, \{5, 6\}\}$
$[\mathcal{P}_3]$	$\{\{1, 2\}, \{3, 4\}, \{5, 6, 7\}\}; \{\{1, 2\}, \{3, 5\}, \{4, 6, 7\}\}; \{\{1, 2\}, \{3, 6\}, \{4, 5, 7\}\};$ $\{\{1, 2\}, \{3, 7\}, \{4, 5, 6\}\}; \{\{1, 3\}, \{2, 4\}, \{5, 6, 7\}\}; \{\{1, 3\}, \{2, 5\}, \{4, 6, 7\}\};$ $\{\{1, 3\}, \{2, 6\}, \{4, 5, 7\}\}; \{\{1, 3\}, \{2, 7\}, \{4, 5, 6\}\}; \{\{2, 3\}, \{1, 4\}, \{5, 6, 7\}\};$ $\{\{2, 3\}, \{1, 5\}, \{4, 6, 7\}\}; \{\{2, 3\}, \{1, 6\}, \{4, 5, 7\}\}; \{\{2, 3\}, \{1, 7\}, \{4, 5, 6\}\}$
$[\mathcal{P}_4]$	$\{\{1\}, \{2, 3\}, \{4, 5, 6, 7\}\}; \{\{2\}, \{1, 3\}, \{4, 5, 6, 7\}\}; \{\{3\}, \{1, 2\}, \{4, 5, 6, 7\}\}$
$[\mathcal{P}_5]$	$\{\{1, 2, 3\}, \{4, 5, 6, 7\}\}$

**Proposition A.6.** *Let  $(N, \lambda, \tau)$  be an individualistic society. A partition is in the Core if and only if it is generated by Algorithm A.5.*

*Proof.* We first show that any partition generated by Algorithm A.5 is in the Core for an individualistic society.

In Step 1 of the algorithm, the bottom individual among the remaining individuals chooses one of the individual's most preferred coalitions. Thus, any partition generated by the algorithm in Step 3 can be obtained by the CotW procedure and is thus in the Core for an individualistic society by Proposition 4.3. The equivalent partitions generated by the algorithm in Step 4 are in the Core for an individualistic society by Proposition A.2.

For the reverse direction, recall that any partition in the Core can be constructed by CotW. Steps 1-3 of the algorithm generate all CotW partitions where the bottom individual among the remaining individuals chooses. Step 4 then includes also all equivalent partitions where any weak individual among the remaining individuals chooses and thereby all partitions that can be reached by CotW, i.e., all Core partitions by Proposition 4.3.  $\square$

The algorithm to find the Core elements in egalitarian societies proceeds along the same lines, but now the top individual among the remaining individuals chooses first.

**Algorithm A.7.** Let  $(N, \lambda, \tau)$  be an egalitarian society. Let  $v = v^0$  with  $v^0$  the initial decision node. The initial set of arcs  $A$  is equal to the empty set.

**Step 1** Define  $n(v) = n - \sum_{a \in A^v} \mu_a$ . If  $n(v) > 0$ , consider all upward adjacent coalitions of  $n - n(v) + 1$ . Denote the number of upward adjacent coalitions which are most preferred by  $n - n(v) + 1$  by  $k$ . Add  $k$  arcs starting at  $v$  and ending at distinct new nodes. The arcs all have different weights, each equal to the size of a most preferred upward adjacent coalition of  $n - n(v) + 1$ . If  $n(v) = 0$ , then  $v$  is a terminal node and no new decision node is created.

**Step 2** For each new decision node created in Step 1, go back to Step 1. Step 2 finishes when there are no new decision nodes anymore.

**Step 3** Consider the weighted rooted directed tree  $(V, A, v^0)$  resulting from Steps 1 and 2. For every terminal node  $v$ , let  $(a_1, \dots, a_\ell)$  be the arcs on the path from  $v^0$  to  $v$ , and generate the partition

$$\left\{ [1, \mu_{a_1}], \dots, \left[ \sum_{k=1}^{\ell-1} \mu_{a_k} + 1, n \right] \right\}.$$

**Step 4** Collect all the partitions created in Step 3 and add all equivalent partitions.

**Proposition A.8.** Let  $(N, \lambda, \tau)$  be an egalitarian society. A partition is in the Core if and only if it is generated by Algorithm A.7.

As Algorithm A.7 is the mirror image of Algorithm A.5, we omit the proof of Proposition A.8.

#### Proof of Theorem 4.4.

The proof consists of three parts. First, we prove non-emptiness of the Core. Second, we prove homophily and finally coalitional assortativity of each Core element.

1. By Propositions A.6 and A.8, all Core elements are obtained by Algorithms A.5 and A.7. Note that in each step of the algorithm a top/bottom individual exists and is matched to a coalition in the partition. Thus, after at most  $n$  steps we reach a partition in the Core, i.e., the Core is non-empty.

2. Towards a contradiction, suppose there is a Core partition with a coalition such that  $|S| > \bar{\rho}$  and there is  $i, j \in S$  such that  $\lambda_i \neq \lambda_j$ . Then, by Lemmas A.3 and A.4 there exists a profitable deviation, a contradiction which establishes homophily.

3. Reconsider the Choice of the Weakest/Strongest procedure. In an egalitarian society, a strongest individual among the remaining individuals chooses a coalition  $S$  which maximizes  $\bar{u}(S)$ . Whenever the remaining number of individuals is greater than or equal to  $\rho$ , maximality can only be achieved if  $|S| - 1$  others with the highest productivities are chosen. Coalitional assortativity follows. For individualistic societies, a weakest individual among the remaining individuals chooses a coalition  $S$  which maximizes  $\tau(|S|)$  and among all such coalitions the one that minimizes  $\bar{u}(S)$  since a higher average payoff decreases the own relative payoff. Whenever the remaining number of individuals is greater than or equal to  $\rho$ , minimality can only be achieved if the other coalition members have the lowest productivity. Coalitional assortativity follows. □

**Proof of Proposition 4.7.** Consider the Choice of the Weakest procedure. Recall that  $\tau$  is strictly increasing up to  $\bar{\rho}$ . Thus, if possible, the weakest remaining individual chooses a coalition with the other  $\bar{\rho} - 1$  remaining individuals with the lowest productivities; this coalition is unique by complete heterogeneity. Otherwise, as  $\tau$  is strictly increasing, the weakest remaining individual chooses a coalition with all remaining individuals. We thus obtain the partition  $[1, z], [z + 1, z + \bar{\rho}], \dots, [n - \bar{\rho} + 1, n]$ , where  $\ell$  and  $z$  are non-negative integers such that  $z < \bar{\rho}$  and the number of individuals in the society equals  $n = \ell\bar{\rho} + z$ . □

**Proof of Proposition 4.9.** Consider the Choice of the Strongest Procedure. Assume that there are at least  $\bar{\rho}$  remaining individuals. We argue that the strongest remaining individual chooses a coalition with the other  $\bar{\rho} - 1$  remaining individuals with the highest productivities; this coalition is unique by complete heterogeneity. It follows from complete heterogeneity that the material payoff for the strongest remaining individual is higher than in any other coalition with  $\bar{\rho}$  members chosen from the remaining individuals. Moreover, since  $\tau(\bar{\rho}) = 1$ , this payoff is also higher than in any coalition of size larger than  $\bar{\rho}$  chosen from the remaining individuals. The most productive coalitions of smaller size can be obtained by consecutively



dropping the bottom individual. It follows from cohesiveness of adjacent groups, that the material payoff of the strongest remaining individual is then lower. If there are less than  $\bar{\rho}$  remaining individuals, then the strongest remaining individual chooses a coalition with all remaining individuals. The most productive coalitions of smaller size can be obtained by consecutively dropping the bottom individual. It follows from cohesiveness of adjacent groups, that the material payoff of the strongest remaining individual is then lower. We thus obtain the partition  $[1, \bar{\rho}], [\bar{\rho} + 1, \dots, 2\bar{\rho}], \dots, [\ell\bar{\rho} + 1, \dots, \ell\bar{\rho} + z]$ , where  $\ell$  and  $z$  are non-negative integers such that  $z < \bar{\rho}$  and the number of individuals in the society equals  $n = \ell\bar{\rho} + z$ .  $\square$

**Proof of Theorem 4.11.** Let  $\ell$  and  $z$  be non-negative integers such that  $z < \bar{\rho}$  and the number of individuals in the society equals  $n = \ell\bar{\rho} + z$ . By the proof of Proposition 4.9, in the egalitarian case, the total material payoff equals  $\sum_{i \in N} \lambda_i - (1 - \tau(z)) \sum_{i=n-z+1}^n \lambda_i$ . For the individualistic society, by the proof of Proposition 4.7, total material payoffs are equal to  $\sum_{i \in N} \lambda_i - (1 - \tau(z)) \sum_{i=1}^z \lambda_i$ .

If  $z = 0$ , i.e.,  $n$  is an integer multiple of  $\bar{\rho}$ , both societies reach the maximal material efficiency equal to  $\sum_{i \in N} \lambda_i$ . If  $z > 0$ , then  $z$  individuals are part of a coalition of size  $z$ , which is not maximally productive. Thus  $1 > \tau(z) > 0$  and by complete heterogeneity, we have  $\sum_{i=1}^z \lambda_i > \sum_{i=n-z+1}^n \lambda_i$ . We conclude that total material payoffs are lower in the individualistic society than in the egalitarian society.  $\square$

**Proof of Theorem 5.2:** Notice that the Core trivially satisfies deterrence of external deviations. In order to show that it coincides with the Myopic Stable Set, we have to show that it satisfies iterated external stability. The minimality requirement follows from the fact that the Core is a subset of the Myopic Stable Set.

### Egalitarian Societies

The proof of iterated external stability is constructive using the following procedure, which, given an initial partition  $\mathcal{P}_0 \in \mathbb{P} \setminus C$ , generates a finite sequence of productive coalitions  $S^1, \dots, S^\ell$  and partitions  $\mathcal{P}_1, \dots, \mathcal{P}_\ell$  such that, for  $k = 1, \dots, \ell$ ,  $\mathcal{P}_k \in \tilde{f}_{S^k}(\mathcal{P}_{k-1})$  and  $\mathcal{P}_\ell$  is a partition in the Core. The coalitions  $S^k$  are all disjoint and chosen such that their members have the highest productivity among agents in  $N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa)$ .

Start with a partition  $\mathcal{P}_0 \in \mathbb{P} \setminus C$ . We define  $k = 1$ .

(i) If  $|N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa| < \rho$ , then the procedure ends. Otherwise, move to (ii).

(ii) Consider a coalition  $T \subseteq N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa$  with  $\bar{u}(T) \in \operatorname{argmax}\{\bar{u}(S) : S \cap \bigcup_{\kappa=1}^{k-1} S^\kappa = \emptyset\}$ . By monotonicity of  $\bar{u}$  in  $\lambda$ , for all  $i \in T$  and for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup T)$  it holds that  $\lambda_i \geq \lambda_j$ . If  $f_T(\mathcal{P}_{k-1}) \neq \emptyset$ , then define  $S^k = T$ . If  $f_T(\mathcal{P}_{k-1}) = \emptyset$ , then some member of

$T$  belongs to a productive coalition  $S \in \mathcal{P}_{k-1}$  such that  $\bar{u}(S) = \bar{u}(T)$ . For all  $i \in S$ , for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup S)$ , it holds that  $\lambda_i \geq \lambda_j$ . In this case, define  $S^k = S$ . In both cases, define  $\mathcal{P}_k = \mathcal{Q}(\mathcal{P}_{k-1}, S^k)$ . Increase  $k$  by 1 and move back to (i).

The iteration takes  $\ell \leq n$  steps and finally generates the partition  $\mathcal{P}_\ell$ . The collection of productive coalitions in  $\mathcal{P}_\ell$  is given by  $\{S^k \mid k \in \{1, \dots, \ell\}\}$ . It holds that  $\mathcal{P}_\ell \in \tilde{f}^\ell(\mathcal{P}_0)$  since, for every  $k \in \{1, \dots, \ell\}$ ,  $\mathcal{P}_k \in f_{S^k}(\mathcal{P}_{k-1})$  or  $\mathcal{P}_k = \mathcal{P}_{k-1}$ .

It remains to be shown that  $\mathcal{P}_\ell$  belongs to the Core. Note that in each step of the procedure, at least one strongest individual among the remaining individuals is added to one of the individual's most preferred coalitions among the remaining individuals. At the point where less than  $\rho$  agents are left, a strongest individual  $i$  among the remaining agents cannot do better than choosing coalition  $S_i(\mathcal{P}_\ell)$ . Thus, we can retrieve the procedure with the Choice of Strongest procedure and the partition  $\mathcal{P}_\ell$  is in the Core.

### Individualistic Societies

The proof of iterated external stability is constructive using the following procedure, which, given an initial partition  $\mathcal{P}_0 \in \mathbb{P}C$ , generates a finite sequence of productive coalitions  $S^1, \dots, S^\ell$  and partitions  $\mathcal{P}_1, \dots, \mathcal{P}_\ell$  such that, for  $k = 1, \dots, \ell$ ,  $\mathcal{P}_k \in \tilde{f}_{S^k}(\mathcal{P}_{k-1})$  and  $\mathcal{P}_\ell$  is a partition in the Core. The coalitions  $S^k$  are all disjoint and chosen such that their members have the lowest productivity among agents in  $N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa)$ .

Start with a partition  $\mathcal{P}_0 \in \mathbb{P}C$ . Parts (i) and (ii) formalize a sequence of dominations where first a productive coalition with weak individuals forms and so forth until a strongest productive coalition forms. We define  $k = 1$ .

(i) If  $|N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa| < \rho$ , then the procedure ends. Otherwise, move to (ii).

(ii) Consider a coalition  $T \subseteq N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa$  with  $|T| = \min\{\bar{\rho}, |N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa|\}$  such that for all  $i \in T$  and for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup T)$  it holds that  $\lambda_i \leq \lambda_j$ . If  $f_T(\mathcal{P}_{k-1}) \neq \emptyset$ , then define  $S^k = T$ . If  $f_T(\mathcal{P}_{k-1}) = \emptyset$ , then there is  $i' \in T$  who belongs to a coalition  $S \in \mathcal{P}_{k-1}$  such that  $m_{i'}^{\text{ind}}(S) = m_{i'}^{\text{ind}}(T)$  and  $r_{i'}^{\text{ind}}(S) \geq r_{i'}^{\text{ind}}(T)$ . Since

$$m_{i'}^{\text{ind}}(S) - \bar{u}(S) = r_{i'}^{\text{ind}}(S) \geq r_{i'}^{\text{ind}}(T) = m_{i'}^{\text{ind}}(T) - \bar{u}(T) = m_{i'}^{\text{ind}}(S) - \bar{u}(T),$$

we have that  $\bar{u}(S) \leq \bar{u}(T)$ . The construction of  $T$  now implies  $\bar{u}(S) = \bar{u}(T)$ . It follows that, for all  $i \in S$ , for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup S)$ ,  $\lambda_i \leq \lambda_j$ . In this case, define  $S^k = S$ . Define  $\mathcal{P}_k = \mathcal{Q}(\mathcal{P}_{k-1}, S^k)$ . Increase  $k$  by 1 and move back to (i).

The procedure takes  $\ell \leq n/\bar{\rho} + 1$  steps and finally generates the partition  $\mathcal{P}_\ell$ . The collection of productive coalitions in  $\mathcal{P}_\ell$  is given by  $\{S^k \mid k \in \{1, \dots, \ell\}\}$ . It holds that  $\mathcal{P}_\ell \in \tilde{f}^\ell(\mathcal{P}_0)$  since, for every  $k \in \{1, \dots, \ell\}$ ,  $\mathcal{P}_k \in f_{S^k}(\mathcal{P}_{k-1})$  or  $\mathcal{P}_k = \mathcal{P}_{k-1}$ .

It remains to be shown that  $\mathcal{P}_\ell$  belongs to the Core. Note that in each step of the procedure, at least one weakest individual among the remaining agents is added to one of the individual's most preferred coalitions among those remaining individuals. At the point where less than  $\rho$  agents are left, a weakest individual  $i$  among the remaining agents cannot do better than choosing coalition  $S_i(\mathcal{P}_\ell)$ . Thus, we can retrieve the process with the Choice of Weakest procedure and the partition  $\mathcal{P}_\ell$  is in the Core.  $\square$

### A.3 An Example of an Empty Core and a Non-Empty Myopic Stable Set

Consider an individualistic society in the setting of Subsection 6.1, where preferences are determined by a convex combination of material and relative payoffs. Let  $n = 4, \lambda_1 = 1, \lambda_2 = 0.99, \lambda_3 = 0.98, \lambda_4 = 0.25, \tau(1) = 0, \tau(2) = 0.8, \tau(3) = \tau(4) = 1$ , and  $\mu = 0.75$ .

**Core** We show that the Core is empty.

Note first that any individual prefers any non-singleton coalition to staying alone as, for any  $S \in \mathcal{N}$  with  $|S| \geq 2$ , for any  $j \in S$ , it holds that

$$p_j(S) = \tau(|S|) \left( \lambda_j - (1 - 0.75) \cdot \frac{\sum_{i \in S} \lambda_i}{|S|} \right) > 0,$$

since  $\tau(|S|) > 0, \lambda_j \geq 0.25$ , and  $\sum_{i \in S} \lambda_i / |S| < 1$ . Thus, a partition in the Core cannot contain more than one singleton coalition.

**Case 1:** A partition with two coalitions of size 2.

We will show that any member of the coalition  $\{1, 2, 3\}$  is better off compared to staying in a coalition of size 2.

Consider a partition with two coalitions of size 2. For any  $i \in \{1, 2, 3\}$ , we can bound the payoff of an individual  $i$  in a coalition  $S$  of size 2 from above by the highest possible payoff of individual 1 in a coalition of size 2:

$$p_i(S) \leq 0.8 \left( 1 - (1 - 0.75) \cdot \frac{1 + 0.25}{2} \right) = 0.675.$$

In the coalition  $\{1, 2, 3\}$ , the payoff of  $i \in \{1, 2, 3\}$  is bounded from below by

$$p_i(\{1, 2, 3\}) \geq p_3(\{1, 2, 3\}) = 0.98 - (1 - 0.75) \cdot \frac{1 + 0.99 + 0.98}{3} = 0.7325 > 0.675.$$

Thus, a partition with two coalitions of size 2 cannot be part of the Core.

**Case 2:** A partition with one coalition of size 3 and one coalition of size 1.

(a) Suppose the coalition of size 3 is  $\{1, 2, 3\}$ . Then coalition  $\{1, 2, 4\}$  can profitably deviate: individual 4 increases its payoff to above zero, while individuals 1 and 2 benefit

from an increase in relative payoffs by forming a coalition with individual 4 rather than individual 3.

(b) Suppose the coalition  $S$  of size 3 contains individual 4. Let  $i$  be the individual in the singleton coalition. Then the coalition  $\{i, 4\}$  can profitably deviate. Individual  $i$  gets a positive payoff and thus improves upon staying single. In the coalition of size 3, individual 4 has a payoff of

$$p_4(S) \leq p_4(\{2, 3, 4\}) = 0.25 - (1 - 0.75) \cdot \frac{0.99 + 0.98 + 0.25}{3} = 0.065.$$

In the coalition  $\{i, 4\}$ , the payoff of individual 4 satisfies

$$p_4(\{i, 4\}) \geq p_4(\{1, 4\}) = 0.8 \left( 0.25 - (1 - 0.75) \cdot \frac{1 + 0.25}{2} \right) = 0.075 > 0.065.$$

**Case 3:** The grand coalition forms.

In this case, the coalition  $\{2, 3, 4\}$  is better off by deviating, as the material payoffs remain the same, but the relative payoffs improve by the exclusion of individual 1.

Thus, there are no partitions which do not allow for profitable deviations, i.e., the Core is empty.

**Myopic Stable Set** Let  $\mathbb{P}_4$  be the set of partitions with one coalition of size 4, i.e.  $\mathbb{P}_4 = \{\{\{1, 2, 3, 4\}\}\}$ ,  $\mathbb{P}_3$  be the set of partitions with one coalition of size 3, so  $\mathbb{P}_3$  has 4 elements, and  $\mathbb{P}_2$  be the set of partitions with two coalitions of size 2, so  $\mathbb{P}_2$  has 3 elements. We define the set of partitions  $M$  as

$$M = \mathbb{P}_4 \cup \mathbb{P}_3 \cup \mathbb{P}_2 \cup \{\{\{2, 4\}, \{1\}, \{3\}\}\} \cup \{\{\{3, 4\}, \{1\}, \{2\}\}\}$$

and demonstrate that the Myopic Stable Set is equal to  $M$ .

**Deterrence of External Deviations:** We first argue that  $M$  satisfies deterrence of external deviations.

Since the payoff of an individual in a singleton coalition is equal to 0, there are no profitable deviations by a single agent. Therefore, there is no deviation from a partition in  $M$  to the partition  $\{\{1\}, \{2\}, \{3\}, \{4\}\}$ .

We now argue that there are no profitable deviations from partitions in  $M$  to the partitions  $\{\{1, 2\}, \{3\}, \{4\}\}$ ,  $\{\{1, 3\}, \{2\}, \{4\}\}$ ,  $\{\{1, 4\}, \{2\}, \{3\}\}$ , and  $\{\{2, 3\}, \{1\}, \{4\}\}$ . Such a partition can only be reached from partitions in  $\mathbb{P}_2$  when the resulting coalition of size 2 deviates or from partitions in  $\mathbb{P}_3$  when the resulting coalition of size 2 deviates and is part of the coalition of size 3 before the deviation.

There are no such profitable deviations from partitions in  $\mathbb{P}_2$ . A coalition partner of individual 4 in a partition in  $\mathbb{P}_2$  cannot profitably deviate with another individual since material payoffs remain the same and relative payoffs would decrease. The only remaining

possibility from a partition in  $\mathbb{P}_2$  is a deviation by coalition  $\{1, 4\}$ . But then the relative payoff of individual 4 would decrease, whereas the material payoff remains the same.

There are also no such profitable deviations from partitions in  $\mathbb{P}_3$ . The partition  $\{\{1, 2\}, \{3\}, \{4\}\}$  can only be reached from the partitions  $\{\{1, 2, 3\}, \{4\}\}$  and  $\{\{1, 2, 4\}, \{3\}\}$  by a deviation by coalition  $\{1, 2\}$ . This would lead to a decrease in material and relative payoffs for individuals 1 and 2, so it is not profitable. The partition  $\{\{1, 3\}, \{2\}, \{4\}\}$  can only be reached from the partitions  $\{\{1, 2, 3\}, \{4\}\}$  and  $\{\{1, 3, 4\}, \{2\}\}$  by a deviation by coalition  $\{1, 3\}$ . This would lead to a decrease in material and relative payoffs for individual 1, so it is not profitable. The partition  $\{\{1, 4\}, \{2\}, \{3\}\}$  can only be reached from the partitions  $\{\{1, 2, 4\}, \{3\}\}$  and  $\{\{1, 3, 4\}, \{2\}\}$  by a deviation by coalition  $\{1, 4\}$ . It holds that  $p_1(\{1, 3, 4\}) > p_1(\{1, 2, 4\}) = 0.813 > 0.675 = p_1(\{1, 4\})$ , so such a deviation is not profitable for individual 1. Finally, the partition  $\{\{2, 3\}, \{1\}, \{4\}\}$  can only be reached from the partitions  $\{\{1, 2, 3\}, \{4\}\}$  and  $\{\{2, 3, 4\}, \{1\}\}$  by a deviation by coalition  $\{2, 3\}$ . Payoffs for individuals 2 and 3 are higher in coalition  $\{2, 3, 4\}$  than in coalition  $\{1, 2, 3\}$ , which were shown to be higher than their payoffs in any coalition of size 2. Such a deviation is therefore not profitable for individuals 2 and 3.

**Iterated External Stability:** The set  $M$  satisfies iterated external stability, since we argued before that any member of the coalition  $\{1, 2, 3\}$  is better off compared to staying in a coalition of size 1 or 2.

**Minimality:** We show that  $M$  satisfies minimality by arguing that, for every  $\mathcal{P} \in M$ ,  $\tilde{f}^{\mathbb{N}}(\mathcal{P}) = M$ , i.e. the set  $M$  is a so-called closed cycle and therefore each element in it has to be part of the Myopic Stable Set by Theorem 3.9 in Demuynck, Herings, Saulle, and Seel (2019).

Consider any partition in  $M$  with no coalition of size 3 or larger. We argued before that the coalition  $\{1, 2, 3\}$  has a profitable deviation. It holds that  $\mathbb{P}_3 \subseteq \tilde{f}(\{\{1, 2, 3\}, \{4\}\})$  since material payoffs would remain the same and relative payoffs would improve after a deviation by a three-agent coalition different from  $\{1, 2, 3\}$ . The same argument plus the observation that individual 4 obtains a positive payoff shows that  $\mathbb{P}_4 \subset f(\{\{1, 2, 3\}, \{4\}\})$ . It clearly holds that  $\{\{2, 3, 4\}, \{1\}\} \in f(\{\{1, 2, 3, 4\}\})$ . From any partition with a coalition of size 3 different from  $\{1, 2, 3\}$ , individual 4 has a profitable deviation by forming a coalition with the unmatched individual, which reaches all partitions with two coalitions of size 2. It holds that  $\{\{2, 4\}, \{1\}, \{3\}\} \in f(\{\{1, 4\}, \{2, 3\}\})$  and  $\{\{3, 4\}, \{1\}, \{2\}\} \in f(\{\{2, 4\}, \{1, 3\}\})$ . It follows that, for every  $\mathcal{P} \in M$ ,  $M \subseteq \tilde{f}^{\mathbb{N}}(\mathcal{P})$ . Since  $M$  satisfies deterrence of external deviations, we obtain that, for every  $\mathcal{P} \in M$ ,  $\tilde{f}^{\mathbb{N}}(\mathcal{P}) = M$ .

P. Jean-Jacques Herings, Department of Economics, Maastricht University, Maastricht, The Netherlands. E-Mail: P.Herings@maastrichtuniversity.nl

Riccardo D. Saulle, Department of Economics and Management, University of Padova, Padua, Italy. E-mail: [riccardo.saulle@unipd.it](mailto:riccardo.saulle@unipd.it)

Christian Seel, Department of Economics, Maastricht University, Maastricht, The Netherlands. E-Mail: [C.Seel@maastrichtuniversity.nl](mailto:C.Seel@maastrichtuniversity.nl)

## References

- ACEMOGLU, D., G. EGOROV, AND K. SONIN (2008), “Coalition Formation in Non-Democracies”, *Review of Economic Studies*, 75, 987–1009.
- AMBEC, S., AND L. EHLERS (2008), “Sharing a River Among Satiabile Agents”, *Games and Economic Behavior*, 64, 35–50.
- BANERJEE, S., H. KONISHI, AND T. SÖNMEZ (2001), “Core in a Simple Coalition Formation Game,” *Social Choice and Welfare*, 18, 135–153.
- BARBERÀ, S., C. BEVIÁ, AND C. PONSATÍ (2015), “Meritocracy, Egalitarianism and the Stability of Majoritarian Organizations,” *Games and Economic Behavior*, 91, 237–257.
- BEVIÁ, C., L. CÓRCHON, AND A. ROMERO-MEDINA (2017), “Relinquishing Power, Exploitation and Political Unemployment in Democratic Organizations,” *Social Choice and Welfare*, 49, 735–753.
- BILBAO, J.M., A. JIMÉNEZ, AND J.J. LÓPEZ (1998), “The Banzhaf Power of Convex Geometries”, *Mathematical Social Sciences*, 36, 157–173.
- BOGOMOLNAIA, A., AND M.O. JACKSON (2002), “The Stability of Hedonic Coalition Structures,” *Games and Economic Behavior*, 38, 201–230.
- BRACHA, A., U. GNEEZY, AND G. LOEWENSTEIN (2015), “Relative Pay and Labor Supply,” *Journal of Labor Economics*, 33, 297–315.
- CARD, D., A. MAS, E. MORETTI, AND E. SAEZ (2012), “Inequality at Work: The Effect of Peer Salaries on Job Satisfaction,” *American Economic Review*, 102, 2981–3003.
- CHWE, M.S.-Y. (1994), “Farsighted Coalitional Stability,” *Journal of Economic Theory*, 63, 299–325.
- CLARK, A.E., P. FRIJTERS, AND M.A. SHIELDS (2008), “Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles,” *Journal of Economic Literature*, 46, 95–144.
- DEMUYNCK, T., P.J.J. HERINGS, R.D. SAULLE, AND C. SEEL (2019), “The Myopic Stable Set for Social Environments,” *Econometrica*, 87, 111–138.
- DUESENBERY, J.S. (1949), *Income, Saving, and the Theory of Consumer Behavior*, Harvard University Press, Cambridge.

- DUFWENBERG, M., P. HEIDHUES, G. KIRCHSTEIGER, F. RIEDEL, AND J. SOBEL (2011), “Other-regarding Preferences in General Equilibrium,” *Review of Economic Studies*, 78, 613–639.
- FARRELL, J., AND S. SCOTCHMER (1988), “Partnerships,” *Quarterly Journal of Economics*, 103, 279–297.
- FERRER-I-CARBONELL, A. (2005), “Income and Well-being: An Empirical Analysis of the Comparison Income Effect,” *Journal of Public Economics*, 89, 997–1019.
- FRANK, R.H. (1985), *Choosing the Right Pond: Human Behavior and the Quest for Status*, Oxford University Press, New York.
- GOYAL, S., P. HERNÁNDEZ, G. MARTÍNEZ-CÁNOVAS, F. MOISAN, M. MUÑOZ-HERRERA, AND A. SÁNCHEZ (2018), “Integration and Diversity,” Working Paper, 1–56.
- GREENBERG, J., AND S. WEBER (1986), “Strong Tiebout Equilibrium under Restricted Preferences Domain,” *Journal of Economic Theory*, 38, 101–117.
- HAFER, C. (2006), “On the Origins of Property Rights: Conflict and Production in the State of Nature,” *Review of Economic Studies*, 73, 119–143.
- HART, S., AND M. KURZ (1983), “Endogenous Formation of Coalitions,” *Econometrica*, 51, 1047–1064.
- HERINGS, P.J.J., A. MAULEON, AND V. VANNETELBOSCH (2009), “Farsightedly Stable Networks,” *Games and Economic Behavior*, 67, 526–541.
- HOBBS, T. (1651), *Leviathan*, Reprinted in Baltimore: Penguin Books, 1968.
- JORDAN, J.S. (2006), “Pillage and Property,” *Journal of Economic Theory*, 131, 26–44.
- JORDAN, J.S., AND D. OBADIA (2015), “Stable Sets in Majority Pillage Games,” *International Journal of Game Theory*, 44, 473–486.
- KETS, W. AND A. SANDRONI (2019), “A Belief-based Theory of Homophily,” *Games and Economic Behavior*, 115, 410–435.
- KÓCZY, L.Á. (2018), *Partition Function Form Games*, Theory and Decision Library C, Volume 48, Springer International Publishing.
- KÓCZY, L.Á., AND L. LAUWERS (2007), “The Minimal Dominant Set Is a Non-empty Core-extension,” *Games and Economic Behavior*, 61, 277–298.



- LUTTMER, E.F.P. (2005), “Neighbors as Negatives: Relative Earnings and Well-being,” *Quarterly Journal of Economics*, 120, 963–1002.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2012), “Social Decision Theory: Choosing within and between Groups,” *Review of Economic Studies*, 79, 1591–1636.
- MILCHTAICH, I., AND E. WINTER (2002), “Stability and Segregation in Group Formation,” *Games and Economic Behavior*, 38, 318–346.
- MORELLI, M., AND I.-U. PARK (2016), “Internal Hierarchy and Stable Coalition Structures,” *Games and Economic Behavior*, 96, 90–96.
- MUTHOO, A. (2004), “A Model of the Origins of Basic Property Rights,” *Games and Economic Behavior*, 49, 288–312.
- NEUMANN, J. VON, AND O. MORGENSTERN (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, New Jersey.
- PAGE, F.H., JR., AND M. WOODERS (2009), “Strategic Basins of Attraction, the Path Dominance Core, and Network Formation Games,” *Games and Economic Behavior*, 66, 462–487.
- PEŃSKI, M., AND B. SZENTES (2013), “Spontaneous Discrimination,” *American Economic Review*, 103, 2412–2436.
- PICCIONE, M., AND R. RAZIN (2009), “Coalition Formation under Power Relations,” *Theoretical Economics*, 4, 1–15.
- PICCIONE, M., AND A. RUBINSTEIN (2007), “Equilibrium in the Jungle,” *Economic Journal*, 117, 883–896.
- RAY, D. (2007), *A Game-theoretic Perspective on Coalition Structures*, Oxford University Press, Oxford.
- RAY, D., AND R. VOHRA (2014), “Coalition Formation,” in H.P. Young and S. Zamir (eds.), *Handbook of Game Theory, Volume 4*, Elsevier, Amsterdam, pp. 239–326.
- RAY, D., AND R. VOHRA (2015), “The Farsighted Stable Set,” *Econometrica*, 83, 977–1011.
- WATTS, A. (2007), “Formation of Segregated and Integrated Groups,” *International Journal of Game Theory*, 35, 505–519.
- WILSON, R.J. (1985), *Introduction to Graph Theory*, Third Edition, Longman Scientific and Technical, New York, New York.