Cooperative and Strategic Games in Network Economics

Der Fakultät für Wirtschaftswissenschaften der Universität Paderborn zur Erlangung des akademischen Grades

Doktors der Wirtschaftswissenschaften

– Doctor rerum politicarum –

vorgelegte Inauguraldissertation von

Nils Röhl

geboren am 8.6.1984 in Koblenz

Paderborn, 31. Oktober 2014

Erstgutachter

Prof. Dr. Claus-Jochen Haake
Fakultät für Wirtschaftswissenschaften
Universität Paderborn

Zweitgutachter

Prof. Dr. Herbert Dawid
Institut für mathematische Wirtschaftsforschung
Universität Bielefeld

Acknowledgments

During the recent years, I greatly benefited from the immense encouragement of numerous friends and colleagues who supported me in writing my thesis.

First of all, I would like to thank my supervisor Claus-Jochen Haake for his permanent guidance and continuous motivation. He always kept track of the progress of my thesis and helped me with many valuable comments. In this respect, I am also deeply indebted to Sonja Brangewitz who read large parts of my work and shared her captivating ideas in numerous fruitful discussions. Further thanks go to many other colleagues in Paderborn and Bielefeld who made my time as a PhD-student an enjoyable and enriching chapter in life. Out of these, I would like to especially mention Jochen Manegold who accompanied me in experiencing quite a number of ups an downs and made sharing an office with him a great pleasure. In addition, I am as well grateful to Herbert David for his support and his effort taken in examining my thesis.

Large parts of this work have been established together with Ana Mauleon, Vincent Vannetelbosch, and Berno Büchel who co-authored the third and the fourth chapter, respectively. Especially Berno helped me in doing my first steps on academic paths and widened my horizon in the course of our collaboration. Moreover, I thank Ana and Vincent for enabling a research stay at the Center for Operations Research and Econometrics (CORE) in Louvain-la-Neuve and for providing hospitality. The stay had been funded by the German Academic Exchange Service (DAAD) and I gratefully acknowledge the financial support.

I moreover appreciate fundings by the German Research Foundation (DFG) within the Collaborative Research Center "On-The-Fly Computing" (SFB 901).

Last but by far not least, special thanks go to my family and my girl-friend for their patience, their unconditional support, and their enduring belief in me. I am especially indebted to my sister Sabine who proof-read large parts of this work and to Claudia who always reminded me that there is something more than my thesis.

Nils Röhl

Contents

List of Figures

Chapter 1

Introduction

It is a well-accepted fact that doings and decisions of virtually every acting subject – whether it is an individual, an organization, or even an entire country – do inevitably influence and connect to the actions of other subjects. Depending on the context, these mutual interdependencies induce network-like relationships which may take various forms. In daily life, for example, friendships and interpersonal relations between people span cohesive networks in which they communicate and exchange information. Especially in recent times, these structures take concrete shapes since more and more people are organizing themselves in social networks like Facebook. But similar considerations also apply on institutional levels. Companies, for instance, form reciprocal agreements which induce a high degree of collaboration between them. This not only impacts competition but it also influences the product variety obtainable in the market. Moreover, many countries enter bilateral or even multilateral contracts which affect trade flows on a global level. Consequently, for many years now, investigating economic networks has become a key element of economic science and the number of publications dealing with this subject is continuously increasing. In particular, the studies presented in this thesis also contribute to this field of research.

Since network structures are characteristic for many settings and applications, the corresponding analysis is divided into several (overlapping) branches. Among them one could mention not only stability and efficiency of network structures but also communication issues or public good provision, to name but a few. In addition to this, many laboratory experiments and econometric studies have been conducted in this context. As these examples already indicate, economic networks have been approached in literature in manifold and highly heterogeneous ways. Thus, it is impossible to embed all of them in a general framework. Instead, the objective of this thesis is to highlight cooperative as well as strategic games in network economics in order to explore this aspect in more detail and to complement it by analyzing issues which have not been covered so far.

Roughly speaking, the vast majority of research questions in network theory that utilize game theoretic tools is addressed in three specific branches. The first one concentrates on allocation rules for network games, i.e., on the issue of cooperatively allocating welfare among the members of a (given) network. The main objective of the second one is, on the other hand, to study strategic network formation in order to predict which networks are likely to occur. Last but not least, the third branch uses game theoretic concepts for analyzing locational competition on networks. As a consequence of this division, the main part of this thesis consists of three self-contained chapters (i.e., each can be read independently of the others), and each of them focuses on one of the aforementioned topics. Chapter 2 looks at allocation rules for network games whereas Chapter 3 concentrates on non-cooperative network formation and Chapter 4 on locational competition on networks. The main goal is to model the economic problems addressed in these branches more realistically in order to enlarge the field of possible applications. This is done by allowing for more flexibility in the formal substructure of the standard approaches. In order to give a more precise idea about the proceeding, each of the branches will be outlined briefly in the remainder of this introduction by surveying the most important contributions and highlighting the main research questions. However, each chapter also contains a comprehensive introduction that leads more into the details of the corresponding topic and discusses interdisciplinary literature as well.

1.1 Allocation Rules for Network Games

The origins of this branch date back to cooperative game theory. Traditionally, cooperative games with transferable utility (TU games) focus on situations where a group of individuals generates some welfare by cooperation. The goal is to find appropriate ways for allocating the welfare among the members of the group. This is, of course, far away from being a trivial issue. Nevertheless, there are many studies addressing this issue and providing a variety of possible solutions (also known as allocation rules if the solution is single-valued) which are usually based on certain stability or fairness requirements (see, e.g., Peleg and Sudholter, 2003).

One of the most known allocation rules is the famous Shapley Value (Shapley, 1953) which is remarkable for at least two reasons: On the one hand, it is surprisingly intuitive and, on the other, it can be characterized by a combination of convincing axioms (see, e.g., Roth, 1988). In fact, the Shapley Value is the unique solution concept which is efficient (i.e., the entire welfare is allocated to the individuals), symmetric (i.e., all individuals are treated equally), additive (i.e., if there are no externalities between two TU games, the corresponding allocations are mutually independent, too), and satisfies the dummy axiom (i.e., individuals who do not contribute receive a payoff of zero).[1] Ever since this solution concept was introduced, it has inspired many researchers who have varied and extended Shapley's idea in several ways. Among them is also Myerson (1977) who was the first author who explicitly took network relations between the individuals into account. He assumed that the lines of cooperation are restricted by these relations and he adjusted the Shapley Value accordingly. In fact, he has shown that there is a unique allocation rule (by now known as the Myerson Value) which, on the one hand, allocates the welfare generated by each component of the network only among the corresponding members and, on the other, guarantees equal bargaining power in terms of that if a link between two individuals is deleted, both are affected equally (he called this feature "fairness"). In Myerson (1980), he generalized this approach to more complex network structures, namely to mathematical hypergraphs which he interpreted as conference structures, and he also extended his solution by translating the axiomatization from his previous work to the more general setting. A further prominent allocation rule which is based on the setup of Myerson (1977) is the well-known Position Value (Borm *et al.*, 1992). The main idea of Borm *et al.* (1992) was to construct a two-stage allocation procedure where first each link obtains a share of the welfare which then, subsequently, is split equally between the two corresponding partners. Thereby the authors assumed that the payoff of each link is determined according to a modification of the Shapley Value. For trees (i.e., for networks without cycles) Borm *et al.* (1992) also provided an axiomatic characterization of the Position Value but they were not able to find one for arbitrary networks. Moreover, in the same year, van den Nouweland *et al.* (1992) extended this allocation rule to conference structures, too.

Although the aforementioned papers included networks into the setting of TU games in a reasonable way, they are still subject to at least one major limitation: The au-

[1] In his seminal contribution Shapley actually used slightly different axioms. But the characterization mentioned here is more prominent (cf. Winter, 2002).

thors incorporated network relations between the individuals just for restricting the possible lines of cooperation, but the induced welfare did not depend on the network directly. This assumption is maintained by many works built on Myerson's model (e.g., Albizuri *et al.*, 2005; Hamiache, 1999; Herings *et al.*, 2008). For an overview see Slikker and van den Nouweland (2001). However, Jackson and Wolinsky (1996) were the first to recognize that in several environments the limitation appears to be implausible. In their seminal contribution, they therefore followed a different approach and considered a model in which the welfare of cooperation directly depends on the network structure. Even though this step is straightforward and relatively small from a technical point of view, it allowed for many interesting economic implications and gave fresh momentum to the analysis of economic networks. In fact, Jackson and Wolinsky (1996) not only extended the Myerson Value and its characterization to the generalized setting but they also showed that using this allocation rule induces a certain type of stability. Especially the latter point inspired many papers dealing with network formation (cf. Section 1.2).

Later, Jackson (2005) challenged some fundamental aspects of the Myerson Value. His main criticism was that it is appropriate only in situations where the network describing the relations between the individuals is fixed, as this allocation rule does not take into account that, for example, certain links might be added. Following this idea, Jackson (2005) introduced alternative solutions and provided characterizations by means of axioms which explicitly allowed for some flexibility in the underlying network. In the same year, Slikker (2005a,b) also extended the Position Value to the setting of Jackson and Wolinsky (1996). Since in this model the networks play a more central role, the main motivation that first each link receives a share of the welfare became even more explicit. In fact, due to the higher degree of flexibility, van den Nouweland and Slikker (2012) were able to provide an characterization of the Position Value which parallels the one of the Shapley Value. More specifically, the Position Value is the unique allocation rule which is efficient (i.e., the entire welfare generated by a given network is allocated among the individuals), additive (i.e., if there are no externalities between two network games, the corresponding allocations do not affect each other as well), link anonymous (i.e., if the welfare only depends on the number of links, the same goes for the payoff of the individuals), and satisfies the superfluous link property (i.e., if a link does not affect the generated welfare, it does not influence the individuals' payoffs either).

Two further prominent solution concepts not mentioned so far are the egalitarian allocation rule and the component-wise egalitarian allocation rule (both were introduced in Jackson and Wolinsky (1996) but the authors named them differently). The motivation of the first one is straightforward: The welfare is split equally among all individuals. Under the component-wise egalitarian allocation rule, on the other hand, the focus is on the components. That is, the welfare generated by each of them is equally distributed only among the corresponding members (given that there are no externalities between the components). Jackson and van den Nouweland (2005) showed that the latter allocation rule allows for interesting stability implications.

The second chapter of this thesis is built on or is related to all the aforementioned publications and reconnects network theory to its foundations in cooperative game theory. In fact, the main motivation is two-fold: First, the model of Jackson and Wolinsky (1996) is extended even further in order to capture a class of applications which is neither covered by network theory nor by cooperative game theory. More specifically, almost all of the models dealing with economic networks are limited by the implicit assumption that cooperation takes place only within pairs of individuals. In real life, however, this appears to be unrealistic in many settings. Indeed, cooperation usually takes place not only within pairs but also within larger groups such as departments of an organization, for example, and generically these groups may overlap. Therefore, Chapter 2 focuses on the more general framework of overlapping group structures which extends not only network games but also TU games.
The second goal is to provide a framework which allows analyzing two-stage allocation procedures like the Position Value in a structured way. In the context of overlapping group structures this can be done particularly tractably by exploiting the existence of a unique dual game where individuals and links interchange roles. Proceeding this way allows not only modeling two stages explicitly but also formulating convincing properties for solution concepts which characterize them. Thereby it will be shown, for example, that the Position Value and iteratively applying the Myerson Value can be characterized by similar axiomatizations.

1.2 Strategic Network Formation

In most of the publications surveyed in the previous section, the authors designed allocation rules for situations where some welfare already had been generated by a given network. However, the authors usually did not discuss under which circumstances which structures will emerge. Since this is of fundamental interest, especially in recent years more and more attention has been devoted to the issue of network formation and it has been examined from numerous angles. An intuitive and straightforward approach is to focus on structures which are stable with respect to certain deviations. In order to give a brief overview, the following passages summarize the most groundbreaking and prominent contributions to this approach.

The first work which explicitly modeled network formation has been provided by Aumann and Myerson (1988) who considered a model of sequential link formation. Given a finite set of individuals and starting with the empty network (i.e., there are no links at all), the individuals were allowed to progressively add links according to a given (random) order. Thereby the authors assumed that it is not possible to delete links after they had been formed and the final payoff was determined by the Myerson Value (cf. Section 1.1). For solving this game, Aumann and Myerson (1988) focused on a refinement of Nash equilibria (Nash, 1951), namely on subgame perfect equilibria (Selten, 1965). In this specific setting, this basically means that no pair of players wants to add a link even if the sequence does not start at the empty but at a different network. In particular, subgame perfectness implies that when deciding whether to form a link or not, the individuals also take into account which network might emerge in the end. Although the authors introduced the setup formally, they did not analyze it extensively but rather discussed particular examples in which their approach induced more plausible results than coalition formation in cooperative game theory. However, since the existence of an ordering which specifies which links may be added to the network cannot be guaranteed generically, there are not many economic problems to which the model has been applied. Some years later, Myerson (1991) took an alternative and more convincing approach. He introduced a link formation process in which the individuals simultaneously announce to whom they want to be connected and if two individuals agree to being connected, a link between them is formed. The final payoff of the individuals is thereby supposed to directly depend on the structure of the resulting network. Although Myerson (1991) only briefly mentioned this link formation process and did not develop it in detail, some attention has been devoted to it in literature. Indeed, the process inspired several authors who adopted and varied it in multiple ways (e.g., Arcaute *et al.*, 2013; Bala and Goyal, 2000; Dutta and Mutuswami, 1997; Gilles *et al.*, 2012).

Since Myerson's setup obviously induces a non-cooperative game (by now known as consent game), it is natural to focus on Nash equilibria for finding stable structures. However, although the basic idea of the approach is quite plausible and appealing, using this stability concept has a significant drawback in a network framework.

Applying Nash equilibria is appropriate only in settings where the focus is on deviations of single players. But in Myerson's game, the consent of both corresponding individuals is needed to establish a new link. A straightforward implication of this coordination problem is that every network is always stable when nobody wants to sever one of her connections (like in the empty network, for example). Thus, in a sense, there are too many stable outcomes. The first stability concept which overcame this inappropriateness and was broadly accepted in literature stems from Jackson and Wolinsky (1996). In fact, their work not only gave fresh momentum to the analysis of allocation rules (cf. Section 1.1) but it also was the initial point of a new stream of literature addressing network formation. The main idea of Jackson and Wolinsky (1996) was to focus on single links, i.e., they considered a network to be "pairwise stable" if no single link is changed anymore. Thereby, they assumed that deleting a connection can be done unilaterally, while establishing a new one needs the consent of both corresponding individuals. Phrased differently, a network is pairwise stable if, on the one hand, nobody benefits from severing one of her links and, on the other, if someone would like to add a link, the corresponding partner would suffer from this. In this context, Jackson and Wolinsky (1996) found among other things that often there might be a tension between efficiency and stability, i.e., in many settings it might be the case that none of the networks which is desirable from a social planner's point of view is stable and vice versa.
Although the existence of pairwise stable networks is an issue generically (see, e.g., Hellman, 2013; Jackson and Watts, 2001), motivated by the simple but convincing idea of the stability concept, many researchers applied the basic setup to a variety of applications. Goyal and Joshi (2003, 2006), for example, used the model for analyzing the collaboration of firms and free-trade agreements between countries. But in contrast to Jackson and Wolinsky (1996) they have shown that in these more specific settings stability and efficiency are generically not mutually exclusive. Further contributions, to name but a few, stem from Calvó-Armengol (2004) and Bramoullé and Kranton (2007) who study, respectively, job contact networks and risk sharing in networks.

However, even though the concept of pairwise stability is appealing in many ways, in some settings it might be ineligible and, thus, it has been varied and refined in several ways. The four most prominent modifications are pairwise stability with side payments, pairwise Nash stability, strong stability, and farsighted stability.
As mentioned above, Jackson and Wolinsky (1996) implicitly assume that a link can

be severed unilaterally, while establishing a new one needs just the consents of both corresponding partners. In particular, since other individuals cannot influence these decisions, cutting off a connection or establishing a new one might cause externalities within the network. Taking this into account, Bloch and Jackson (2007) considered a model variation in which side payments between the individuals are possible. In fact, including compensations into the model allows to some extent to internalize the externalities and to decrease the tension between efficiency and stability.

Another major criticism of pairwise stability is that this concept just focuses on alterations of single links. But in some situations, individuals might have an incentive to change several connections at the same time. This aspect has been captured at least partially by Calvó-Armengol and İlkılıç (2009) who combined the approaches of Myerson (1991) and Jackson and Wolinsky (1996). In their paper, the authors considered a network to be "pairwise Nash stable" if and only if it is pairwise stable from the perspective of Jackson and Wolinsky (1996) and Nash stable from the perspective of Myerson (1991). This concept is appropriate in situations where the individuals are able to unilaterally delete several of their connections at the same time and where, if two individuals benefit from forming a link, there occurs no coordination problem as in Myerson (1991). Hence, it is straightforward to show that pairwise stability and pairwise Nash stability coincide if and only if in every pairwise stable network, no player benefits from severing any set of her connections.

Some years earlier already, Jackson and van den Nouweland (2005) chose an even more general approach and examined the "strong stability" of networks. The main idea is that the members of a network might be able to coordinate themselves within subgroups in order to conduct multiple changes at the same time. Consequently, Jackson and van den Nouweland (2005) considered a network to be strongly stable if there exists no subgroup of individuals where all of them could improve by changing the network in such a way that they do not need the consent of individuals who are not contained in the subgroup. In this context, the authors analyzed requirements under which stable and socially optimal outcomes coincide, that is, they discussed conditions guaranteeing that there is no tension between stability and efficiency.

Page *et al.* (2005) challenged a further shortcoming of pairwise stability. In fact, the concept requires that the individuals are not forward-looking, meaning that they only care about the immediate loss (or benefit) of changing a single link. Implicitly, they are supposed to not take into account that this might cause further changes of the network which finally might generate a higher (or lower) benefit. By adopting the concepts of Chwe (1994) to network theory, Page *et al.* (2005) tried to overcome this shortcoming and discussed the "farsighted stability" of networks. The idea is, basically speaking, that the individuals are able to predict which structure might occur at the end of an arbitrarily long sequence of changes. Although Page *et al.* (2005) did not define farsighted stability explicitly, according to their main motivation a network is farsightedly stable if each deviation might lead to a sequence of changes which finally makes the deviating individual(s) worse off. Note that this is obviously closely related to strong stability. The main difference is that here, the individuals behave less cooperatively. They do not coordinate directly but they conduct non-cooperatively those changes which lead to further deviations that finally provide them with a higher benefit. For further publications addressing farsighted stability, see Page and Wooders (2009) or Herings *et al.* (2009), to name but a few.

Another stream of literature which should not be neglected in this introduction deals with dynamic network formation. The first publication addressing this issue is from Bala and Goyal (2000). For specifically chosen payoff functions, the authors studied the best response dynamic in the framework of Myerson's consent game. More precisely, they analyzed a repeated version of the game where the individuals are supposed to play a best response against the strategies chosen in the previous period. In addition to this, Bala and Goyal (2000) assumed that there is a certain probability that the individuals stick to their previous strategy in order to escape perpetual miscoordinations. Since these dynamics induce a Markov chain, the authors focused on the question to which networks the process converges to. To this end, they conducted several simulations and for small numbers of players, the outcomes mostly coincided with the strict Nash equilibria of the one-shot game.
One year later, Watts (2001) followed a slightly different approach by considering a random process in which in each period only one single link is picked randomly with uniform probability. If the link is already contained in the network, then either of the involved individuals is allowed to delete it. If the link is not contained in the network, then the two corresponding individuals may decide to establish the connection (both need to agree) and, at the same time, each of them may sever any of her other connections. Thereby, the individuals are assumed to care only about the immediate benefit, i.e., they are supposed to not be forward-looking. The main insight of the paper was that the process converges to a socially optimal outcome only under quite specific conditions.
In Jackson and Watts (2002), the authors modeled dynamic network formation in a

similar way as Watts (2001) did. They also considered a random process in which in each round exactly one link is picked randomly. However, there are two main differences to Watts (2001). First, inspired by pairwise stability the authors assume that only the selected connection may be altered but the other links remain the same. Second, after the involved individuals have decided whether to alter the connection or not, with some small probability there might occur a mutation which reverses the decision. That is, if the link is now contained in the network it is deleted and vice versa. Given this Markov chain, Jackson and Watts (2002) considered a network to be stochastically stable if it is in the support of the process as the probability of the mutation converges to zero. In fact, they also provide a characterization of these networks (by using "minimal resistance trees") but this characterization is relatively unintuitive and the economic interpretation is not obvious. Taking up this subject, Tercieux and Vannetelbosch (2006) further elaborated on the setting of Jackson and Watts (2002). In doing so, they characterized stochastically stable networks by refining pairwise stability.
Another publication contributing to this stream of literature is Feri (2007). The author analyzed stochastic stability in the framework of Bala and Goyal (2000) by using a slightly varied payoff scheme but analyzing a random process which is similar to the one from Jackson and Watts (2002). Thereby, he found that the set of strict Nash networks from Bala and Goyal (2000) and the set of stochastically stable networks almost coincide.

Similar to the aforementioned literature, the third chapter of this thesis also concentrates on network formation. It is a joint work with Ana Mauleon and Vincent Vannetelbosch, whom I met when I was visiting the Center for Operations Research and Econometrics (CORE) in Louvain-la-Neuve for six months. During my stay we jointly worked out our research question in many fruitful and interesting discussions and developed our model accordingly. Indeed, it complements the aforementioned literature in at least two ways. First, it considers a more general notion of networks. Our main objective is to analyze the formation of group structures where individuals are allowed to engage in several groups at the same time. Formally this means that each link in a network may contain not only two partners but an arbitrary number of individuals. Second, the more distinguishing feature is the formal introduction of constitutions. Each group or link is supposed to have a constitution governing which members may join or leave it. Given these constitutional rules, a network is considered to be stable if none of the groups is altered any more. This, in particu-

lar, implies that the stability of a network depends on explicitly formalized rules of network formation, and the analysis conducted in Chapter 3 is therefore two-fold: It not only focuses on the issue of whether stable networks actually exist but also on how the constitutions need to be designed in order to guarantee stability if the individuals follow a "trial-and-error strategy".
After having constructed the formal substructure, I undertook the analysis of the model in order to derive the results presented in Chapter 3. The trial-and-error behavior, for example, is formalized by means of a random process which is similar to the one of Jackson and Watts (2002). Based on these preliminaries, we further elaborated on the central theme in order to emphasize the main insights of our model. For instance, we show that enhancing the blocking power of the individuals does not necessarily lead to more stability and that a stable network is obtained for sure if and only if there is a certain degree of consent about which feasible deviations (according to the constitutions) are beneficial and which are not. Furthermore, by embedding many-to-many matchings into our setting, we apply the model to job markets with labor unions. To some extent the unions may provide job guarantees and, thus, have influence on the stability of the job market.

1.3 Locational Competition on Networks

The large field of locational competition dates back to the pioneering work of Hotelling (1929). He illustrated the competition between two firms operating in a heterogeneous market by means of a simple but intriguingly intuitive two-stage model. At the first stage, both competitors simultaneously choose a position in the market which Hotelling (1929) modeled as a linear interval. The consumers are supposed to be uniformly distributed along the interval and their utility is linearly decreasing in the distance to the firm they buy from. At the second stage, each competitor chooses a price for her good where the marginal costs are assumed to be constant. Taking the prices and the locations into account, the consumers buy the product which gives a higher benefit to them. Hotelling (1929) thereby assumed that the total demand is totally inelastic. Given this setup, he found what has by now become famous as the principle of minimal differentiation: In the unique subgame perfect Nash equilibrium (cf. Section 1.2) both firms cluster in the middle of the market and choose identical prices. By translating this result to daily life, Hotelling (1929) tried to explain the increasing amount of standardization:

> "Buyers are confronted everywhere with an excessive sameness. [...] there is an incentive to make the new product very much like the old, applying some slight change which will seem an improvement to as many buyers as possible without ever going far in this direction. [...] So general is this tendency that it appears in the most diverse fields of competitive activity, even quite apart from what is called economic life. In politics it is strikingly exemplified." (Hotelling, 1929, p. 54)

Due to its clear and convincing message, the model received a great deal of attention. However, by showing that there actually exists no subgame perfect Nash equilibrium in the model, d'Aspremont *et al.* (1979) proved about 50 years later that Hotelling (1929) was wrong. Furthermore, the authors also reinforced the main message of several other studies which had shown that Hotelling's setup is extremely sensitive with respect to the underlying assumptions (e.g., Downs, 1957; Eaton and Lipsey, 1975; Lerner and Singer, 1937). Even if the fundamentals of the model are relaxed or varied only slightly, it might change the final outcome considerably. These insights lead to a broad discussion about to which extent minimal differentiation is caused by spatial competition since other publications like de Palma *et al.* (1985, 1990) or Rhee (1996), to name but a few, demonstrated that Hotelling's main result can be restored under certain conditions. Indeed, up to now there is no final answer to this question and the discussion is not over yet (e.g., Hehenkamp and Wambach, 2010; Irmen and Thisse, 1998; Król, 2012; Meagher and Zauner, 2004).

The first three works which used networks for modeling the underlying market were Hakimi (1964), Slater (1977) and Wendell and McKelvey (1981).[2] Actually, the main motivation of Hakimi (1964) was to study the problem of finding the best position of a single facility in a network (with respect to certain requirements) but he did not study a competitive environment. However, he briefly mentioned that this would be a reasonable extension. Following this idea, Slater (1977) deepened these considerations. Since he developed his model independently from Hotelling (1929) and as he was only interested in the location choices of the two competitors, he abstracted from the second stage, i.e., he did not take price competition into account. Moreover, he assumed that the two firms enter the market sequentially in order to guarantee the existence of an equilibrium outcome. Nevertheless, although Slater's motivation was different to Hotelling's, for the special class of tree networks he also

[2] In the context of spatial competition, networks are often denoted as graphs. This convention is maintained in Chapter 4 as well.

found that the firms minimally differentiate at the median (which he called centroid) of the tree. For arbitrary networks he was only able to solve the model partially. Independently of the two aforementioned works but motivated by Hotelling (1929) and voting theory (e.g., Downs, 1957), Wendell and McKelvey (1981) tried to solve the problem of finding a "Condorcet-winner" in a network. This issue can be shown to be equivalent to finding a Nash equilibrium if there are two competitors. Since the authors focused on voting theory, similarly to Slater (1977) they did not consider a second stage with price competition. Indeed, they also found that given a tree network, the median is the only Condorcet-winner.

Based on these works, the analysis of locational competition on networks continued in several directions. Hakimi (1983) extended the model of Slater (1977) by constructing a more general setup in which both competitors were allowed to place more than only one facility, i.e., they were allowed to occupy more than only one position (but the number of positions was exogenously fixed). Analogously to Slater (1977), he also focused on the location choices of the competitors and assumed that they enter the market sequentially. This, in particular, implies that there always exists equilibrium outcomes. However, Hakimi (1983) proved that, except for special cases, finding these equilibria is very complex since the optimization problem of the leader as well as of the follower generically is NP-hard and cannot be solved in polynomial time. That is, although there exists an optimal solution for sure, finding it takes much effort. These insights gave birth to a stream of literature which tries to find efficient algorithms for solving the aforementioned issues or variations of them (e.g., Hansen and Labbé, 1988; Kress and Pesch, 2012; Spoerhase and Wirth, 2009). But the vast majority of these contributions did not address the issue of minimal differentiation.

While the studies mentioned in the previous passage focused on situations where the competitors enter the market sequentially, a further stream of literature considered, in contrast to this, the case of simultaneous entry like Wendell and McKelvey (1981) did. A work which is particularly in spirit with Hotelling (1929) stems from Eiselt (1992). He considered a two-stage competition where the competitors first occupy a position in a tree and then, at the second stage, choose their prices. This is, on the one hand, obviously the most straightforward extension of Hotelling's model to a network setup. But, on the other hand, a consequence of proceeding that way is that generically there exists no equilibrium. In order to avoid the non-existence problem, Eiselt (1992) also analyzed a setup without price competition, i.e, he analyzed the case where prices of both competitors are fixed. This had already been discussed by Eiselt and Laporte (1991). As a by-product they thereby confirmed the results of Slater (1977) and Wendell and McKelvey (1981): Given that prices are identical, both competitors will minimally differentiate at the median. Later, Eiselt and Laporte (1993) also analyzed the locational competition between three players on a tree and derived similar results. However, if the analysis is not restricted to trees, abstracting from price competition is not sufficient for guaranteeing the existence of equilibria as has been shown by Knoblauch (1991) and Dürr and Thang (2007). Moreover, even if equilibria exist, finding them is NP-hard. Therefore, most authors who focused only on the first stage, i.e, only on locational competition, needed to impose further requirements in order to generate significant results. For example, de Palma *et al.* (1989) studied a framework in which the individuals' utility not only depends on the distance to the closest competitor but also randomly on other characteristics not captured by the underlying network. Furthermore, Eiselt and Bhadury (1998) elaborated on the existence of Nash equilibria if the competitors have fixed but unequal prices and Gur *et al.* (2012) concentrated on particular subclasses of networks like cacti, to name but a few.

In addition to the aforementioned publications, there are several other works which do not abandon the second stage completely but vary it by considering other forms of price competition or Cournot competition. For example, Lederer and Thisse (1990) and Dorta-González *et al.* (2005) assumed "delivery pricing" instead of "mill pricing". This means, roughly speaking, that at the second stage, the firms do not charge a single price which is the same for every buyer but instead they determine specific prices for each of the nodes of the network. In fact, given some further relatively mild assumptions, this not only guarantees the existence of equilibrium outcomes but it also assures that these outcomes are socially optimal. In contrast to having price competition at the second stage, some further works studied Cournot competition instead (e.g., Labbé and Hakimi, 1991; Sarkar *et al.*, 1997). More specifically, the firms are again supposed to choose their positions first but then, subsequently, the basic assumption is that they compete in quantities but not in prices. Most of the corresponding studies focus on providing sufficient conditions for assuring the existence of equilibria since those do not necessarily exist in this model variation.

Although the fourth chapter of this thesis (which is a joint work with Berno Büchel where both of us contributed equally) is about locational competition on networks, the main question that is addressed is not restricted to networks and should also be interesting in a more general context. Indeed, our main motivation is to challenge a fundamental aspect of Hotelling's approach. Although the literature on locational competition (not only with respect to networks) is rich and highly diversified, it has been virtually always assumed that individuals who prefer the same position fully agree upon the ranking of the other alternatives, i.e., they have identical preferences or utility functions. This is, of course, hard to justify since it might well be that in real life there are consumers with the same favorite brand but who disagree about the ordering of two other brands. Therefore, in our study we scrutinize whether given outcomes of locational competition rely on the questionable homogeneity requirement or not. If it can be shown, that this assumption is not driving the results, then the model is put on a solid foundation (at least with respect to this crucial aspect).

To fix ideas: Given an underlying network, we consider an equilibrium outcome to be "robust" if it does not depend on the aforementioned homogeneity requirement. A key result of our analysis is the characterization of robust equilibria by four conditions which are based on partitioning the underlying space into hinterlands and competitive zones. Applying this result allows us first of all to judge which of the standard results are robust. In fact, we find that several outcomes do not depend on the homogeneity requirement, but some do. Furthermore, by discussing whether the classical observations of minimal differentiation is a robust phenomenon, we find strong support for an old conjecture that in equilibrium firms form local clusters.

References

Albizuri, M. J., Aurrecoechea, J. and Zarzuelo, J. (2005). Configuration values: Extensions of the coalitional Owen value. *Games and Economic Behavior*, **57**, 1–17.

Arcaute, E., Dyagilev, K., Johari, R. and Mannor, S. (2013). Dynamics in tree formation games. *Games and Economic Behavior*, **79**, 1–29.

Aumann, R. J. and Myerson, R. B. (1988). Endogenous formation of links between players and of coalitions: an application of the shapley value. In Roth (1988), pp. 175–191.

Bala, V. and Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, **68** (5), 1181–1229.

Bloch, F. and Jackson, M. O. (2007). The formation of networks with transfers among players. *Journal of Economic Theory*, **133**, 83–110.

Borm, P., Owen, G. and Tijs, S. (1992). On the position value for communication situations. *SIAM Journal on Discrete Mathematics*, **5** (3), 305–320.

Bramoullé, Y. and Kranton, R. (2007). Risk-sharing networks. *Journal of Economic Behavior & Organization*, **64**, 275–294.

Calvó-Armengol, A. (2004). Job contact networks. *Journal of Economic Theory*, **115**, 191–206.

— and İlkiliç, R. (2009). Pairwise-stability and nash equilibria in network formation. *International Journal of Game Theory*, **38**, 51–79.

Chwe, M. S.-Y. (1994). Farsighted coalitional stability. *Journal of Economic Theory*, **63** (2), 299–325.

d'Aspremont, C., Gabszewicz, J. J. and Thisse, J.-F. (1979). On hotelling's "stability in competition". *Econometrica*, **47** (5), 1145–50.

de Palma, A., Ginsburgh, V., Labbeé, M. and Thisse, J.-F. (1989). Competitive location with random utilities. *Transportation Science*, **23** (4), 244–252.

—, —, Papageorgiou, V. V. and Thisse, J.-F. (1985). The principle of minimum differentiation holds under sufficient heterogeneity. *Econometrica*, **53** (4), 767–781.

—, Hong, G.-S. and Thisse, J.-F. (1990). Equilibria in multi-party competition under uncertainty. *Social Choice and Welfare*, **7** (3), 247–259.

Dorta-González, P., Santos-Peñate, D. R. and Suárez-Vega, R. (2005). Spatial competition in networks under delivered pricing. *Papers in Regional Science*, **84** (2), 271–280.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, **65** (2), 135–150.

Dürr, C. and Thang, N. K. (2007). Nash equilibria in voronoi games on graphs. In L. Arge, M. Hoffmann and E. Welzl (eds.), *Algorithms - ESA 2007*, *Lecture Notes in Computer Science*, vol. 4698, Springer, pp. 17–28.

DUTTA, B. and MUTUSWAMI, S. (1997). Stable networks. *Journal of Economic Theory*, **76**, 322–344.

EATON, B. C. and LIPSEY, R. G. (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. *The Review of Economic Studies*, **42** (1), 27–49.

EISELT, H. A. (1992). Hotelling's duopoly on a tree. *Annals of Operations Research*, **40** (1), 195–207.

— and BHADURY, J. (1998). Reachability of locational nash equilibria. *OR Spektrum*, **20**, 101–107.

— and LAPORTE, G. (1991). Locational equilibrium of two facilities on a tree. *RAIRO*, **25** (1), 5–18.

— and — (1993). The existence of equilibria in the 3-facility hotelling model in a tree. *Transportation Science*, **27** (1), 39–43.

FERI, F. (2007). Stochastic stability in networks with decay. *Journal of Economic Theory*, **135**, 442–457.

GILLES, R. P., CHAKRABARTI, A. and SARANGI, S. (2012). Nash equilibria of network formation games under consent. *Mathematical Social Sciences*, **64**, 159–165.

GOYAL, S. and JOSHI, S. (2003). Networks of collaboration in oligopoly. *Games and Economic Behavior*, **43**, 57–85.

— and — (2006). Bilateralism and free trade. *International Economic Review*, **47** (3), 749–778.

GUR, Y., SABAN, D. and STIER-MOSES, N. E. (2012). The competitive facility location problem in a duopoly, mimeo.

HAKIMI, S. L. (1964). Optimum locations of switching centers and the absolute centers and medians of a graph. *Operations Research*, **12** (3), 450–459.

— (1983). On locating new facilities in a competitive environment. *European Journal of Operational Research*, **12** (1), 29–35.

Hamiache, G. (1999). A value with incomplete communication. *Games and Economic Behavior*, **26**, 59–78.

Hansen, P. and Labbé, M. (1988). Algorithms for voting and competitive location on a network. *Transportation Science*, **22**, 278–288.

Hehenkamp, B. and Wambach, A. (2010). Survival at the center–the stability of minimum differentiation. *Journal of Economic Behavior & Organization*, **76** (3), 853–858.

Hellman, T. (2013). On the existence and uniqueness of pairwise stable networks. *International Journal of Game Theory*, **42** (1), 211–237.

Herings, P. J.-J., Mauleon, A. and Vannetelbosch, V. (2009). Farsightedly stable networks. *Games and Economic Behavior*, **67** (2), 526–541.

—, van der Laan, G. and Talman, D. (2008). The average tree solution for cycle-free graph games. *Games and Economic Behavior*, **62**, 77–92.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, **39** (153), 41–57.

Irmen, A. and Thisse, J.-F. (1998). Competition in multi-characteristics spaces: Hotelling was almost right. *Journal of Economic Theory*, **78**, 76–102.

Jackson, M. O. (2005). Allocation rules for network games. *Games and Economic Behavior*, **51**, 128–154.

— and van den Nouweland, A. (2005). Strongly stable networks. *Games and Economic Behavior*, **51**, 420–444.

— and Watts, A. (2001). The existence of pairwise stable networks. *Seoul Journal of Economics*, **14** (3), 299–322.

— and — (2002). The evolution of social and economic networks. *Journal of Economic Theory*, **106**, 265–295.

— and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, **71** (0108), 44–74.

Knoblauch, V. (1991). Generalizing location games to a graph. *The Journal of Industrial Economics*, **39** (6), 683–688.

Kress, D. and Pesch, E. (2012). Sequential competitive location on networks. *European Journal of Operational Research*, **217** (3), 483–499.

Król, M. (2012). Product differentiation decisions under ambiguous consumer demand and pessimistic expectations. *International Journal of Industrial Organization*, **to appear**.

Labbé, M. and Hakimi, S. L. (1991). Market and locational equilibrium for two competitors. *Operations Research*, **39** (5), 749–756.

Lederer, P. J. and Thisse, J.-F. (1990). Competitive location on networks under delivered pricing. *Operations Research Letters*, **9**, 147–153.

Lerner, A. P. and Singer, H. W. (1937). Some notes on duopoly and spatial competition. *The Journal of Political Economy*, **45** (2), 145–186.

Meagher, K. J. and Zauner, K. G. (2004). Product differentiation and location decisions under demand uncertainty. *Journal of Economic Theory*, **117** (2), 201–216.

Myerson, R. B. (1977). Maximin facility location. *Journal of Research*, **79** (B), 107–115.

— (1980). Conference structures and fair allocation rules. *International Journal of Game Theory*, **9** (3), 169–182.

— (1991). *Game Theory: Analysis of conflict*. Cambridge University Press.

Nash, J. (1951). Non-cooperative games. *Annals of Mathematics*, **54** (2), 286–295.

Page, F. H. and Wooders, M. H. (2009). Strategic basins of attraction, the path dominance core, and network formation games. *Games and Economic Behavior*, **66**, 462–487.

—, — and Kamat, S. (2005). Networks and farsighted stability. *Journal of Economic Theory*, **120** (2), 257–269.

Peleg, B. and Sudholter, P. (2003). *Introduction to the Theory of Cooperative Games*. Kluwer Academic Publishers,.

Rhee, B.-D. (1996). Consumer heterogeneity and strategic quality decisions. *Management Science*, **42** (2), 157–172.

ROTH, A. (ed.) (1988). *The Shapley Value*, Cambridge: Cambridge University Press.

SARKAR, J., GUPTA, B. and PAL, D. (1997). Location equilibrium for cournot oligopoly in spatially separated markets. *Journal of Regional Science*, **37** (2), 195–212.

SELTEN, R. C. (1965). Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für die gesamte Staatswissenschaft*, **121** (2), 301–324.

SHAPLEY, L. S. (1953). A value for n-person games. In H. W. Khun and A. W. Tucker (eds.), *Contributions to the Theory of Games*, Princeton: Princeton University Press, pp. 307–314.

SLATER, P. J. (1977). Graphs and cooperation in games. *Mathematics of Operations Research*, **2** (3), 225–229.

SLIKKER, M. (2005a). A characterization of the position value. *International Journal of Game Theory*, **33**, 505–514.

— (2005b). Link monotonic allocation schemes. *International Game Theory Review*, **7** (4), 473–489.

— and VAN DEN NOUWELAND, A. (2001). *Social and Economic Networks in Cooperative Game Theory*. Kluwer Academic Publishers,.

SPOERHASE, J. and WIRTH, H.-C. (2009). (r, p)-centroid problems on paths and trees. *Theoretical Computer Science*, **410**, 5128–5137.

TERCIEUX, O. and VANNETELBOSCH, V. (2006). A characterization of stochastically stable networks. *International Journal of Game Theory*, **34**, 351–369.

VAN DEN NOUWELAND, A., BORM, P. and TIJS, S. (1992). Allocation rules for hypergraph communication situations. *International Journal of Game Theory*, **20** (3), 255–268.

— and SLIKKER, M. (2012). An axiomatic characterization of the position value for network situations. *Mathematical Social Sciences*, **Forthcoming**.

WATTS, A. (2001). A dynamic model of network formation. *Games and Economic Behavior*, **34**, 331–341.

WENDELL, R. E. and MCKELVEY, R. D. (1981). New perspectives in competitive location theory. *European Journal of Operational Research*, **6** (2), 174–182.

WINTER, E. (2002). The shapley value. In R. J. Aumann and S. Hart (eds.), *The Handbook of Game Theory with Economic Applications*, Amsterdam: North-Holland, vol. 3, pp. 2025–2054.

Chapter 2

Two-Stage Rules

2.1 Introduction

Ever since the beginnings of economic theory, finding appropriate as well as convincing ways for distributing welfare has been one of the fundamental problems. In the context of cooperation between individuals, there are two branches, namely the Theory of Cooperative Games and, more recently, the Theory of Economic Networks, that especially focus on this question and propose several solutions, which are called values or allocation rules. This study relates to both approaches. More specifically, the main purpose is to formalize and examine two-stage allocation procedures. To gain an idea of this issue, consider situations where individuals cooperate in several groups or coalitions which may overlap, like in departments of organizations, for example. The main motivation of two-stage allocation procedures is that the generated welfare is not distributed to the individuals directly but first to the groups and then, in a second step, within each group. In fact, for modeling this is it necessary to extend the aforementioned approaches, Network Theory and Cooperative Game Theory, to a more general model.

To understand the difference between Network Theory and Cooperative Game Theory it is necessary to analyze the view of cooperation in both branches. Cooperative games with transferable utility (TU games) provide a theoretical framework for modeling environments where the individuals are partitioned in coalition structures, i.e., some individuals act together in groups to achieve a common goal but standardly it is assumed that nobody is a member of different groups at once (see, e.g., Aumann and Drèze, 1974). In Network Theory, on the other hand, coalitions are replaced by networks which represent bilateral relationships between individuals (see, e.g., Jackson, 2008). Indeed, one of the most characteristic features of networks is that they exhibit some overlapping structure. However, in these models, cooperation generically takes place only within pairs of individuals, which is a limitation as well. Consider, for example, the models from Radner (1993), Bolton and Dewatriopont (1994), or Arenas *et al.* (2010). Here, the authors use networks for representing the structure of organizations. Within each network, the information is shared along the links: that is, the information is exchanged among pairs of individuals. In general, however, employees also work together within larger groups, such as projects, departments, or divisions, and in many environments these groups may overlap, which neither could be depicted by using bilateral networks nor by using TU games.

A convenient tool for modeling overlapping coalition structures are mathematical hypergraphs which can be interpreted as a kind of overlapping coalition structure. Up to now, there is no general notion for these structures. Depending on the context they are called "conference structures" (e.g., Myerson, 1980), "affiliation networks" (e.g., Wasserman and Faust, 1994), "overlapping coalitions" (e.g., Chalkiadakis *et al.*, 2010), "clubs" (e.g., Fershtman and Persitz, 2012) or simply "hypergraphs" (e.g., Jorzik, 2012; van den Nouweland *et al.*, 1992). To stress both, the origins from Network Theory as well as from Cooperative Game Theory, in this paper they will be called coalitional networks.

Formalizing two-stage allocation procedures is in line with several other works from literature (e.g., Aumann and Drèze, 1974; Borm *et al.*, 1992; Hart and Kurz, 1983; Owen, 1977). More specifically, in their publications the authors introduce and discuss particular examples, like the Aumann-Drèze Value, Owen's Value or the Position Value. Another contribution stems from Winter (1989) who analyzes cooperative games with exogenously-given hierarchical structures. In this context, he extends the Aumann-Drèze Value and Owen's Value to arbitrary numbers of stages. However, these publications do not provide a general framework for discussing two-stage allocation procedures in a structured way. In the context of coalitional networks, this can be done tractably by exploiting the existence of a unique dual network which can be interpreted as some kind of dual game where individuals and coalitions (or links, respectively) interchange roles. This allows not only modeling two stages explicitly but also formulating convincing properties for solution concepts which characterize them. Indeed, it is possible to integrate the aforementioned examples into the extended setting presented here. As it shall be proven in this study, these solution concepts can be characterized by using similar axiomatizations. In particular, the second part of the paper is strongly motivated by van den Brink (2007). In his work the author shows that the Shapley Value (Shapley, 1953) and Equal Division Solutions can be described by means of similar axioms. This leitmotiv will be found twice in the present study: Not only by showing that the aforementioned two-stage allocation procedures satisfy similar axioms, but also by extending the characterizations from van den Brink (2007) to coalitional networks.

Some further publications should be mentioned in the context of this study. One of the first authors theorizing overlapping coalition structures was Myerson (1980). His study is motivated by TU games and the focus is on conference structures which are modeled by means of hypergraphs. However, Myerson's approach differs from the present work in at least one important aspect. He uses overlapping coalition structures only for restricting the possible lines of cooperation between the individuals, but the generated value or welfare, respectively, does not depend on the underlying communication structure directly. This assumption is maintained in most of the works built on Myerson's publication (e.g., Albizuri *et al.*, 2005; van den Nouweland *et al.*, 1992). In recent years, computer scientists also tended to pay more and more attention to Cooperative Game Theory. They use TU games to model the behavior of autonomous agents, and in this context they also investigate overlapping coalition structures. However, most of the corresponding publications (e.g., Chalkiadakis *et al.*, 2010) focus on extensions of the core which, will not be discussed here.
Although several notions and concepts used in this work are adopted from Network Theory (e.g., from Jackson, 2005; Jackson and Wolinsky, 1996), overlapping coalition structures have only been attempted by few scholars in this field. Among them are Fershtman and Persitz (2012) who extend the Connections Model to the generalized framework and Jorzik (2012) who examines allocation rules for hypergraphs. A further interesting work stems from Caulier *et al.* (2012) who study network formation for situations where the set of individuals is already partitioned into coalitions. In this context, the authors examine the stability and efficiency of networks.
Moreover, from a technical point of view the model introduced here also relates to many-to-many matchings (see, e.g., Sotomayor, 2004). In fact, there exists a canonical bijection between many-to-many matchings and coalitional networks. However, virtually all publications from this field do not address two-stage allocation rules but have different objectives.

The remainder of the paper proceeds as follows: In the next section, the model will be introduced formally. This includes the definition of coalitional networks as well as the extension of value functions and characteristic functions to the framework analyzed here. The objective of Section 2.3 is to formalize allocation rules and to characterize two-stage allocation procedures in general. Section 2.4 is devoted to particular examples, like the Position Value, for instance. Finally, Section 2.5 briefly summarizes and presents an agenda for further research.

2.2 The Model

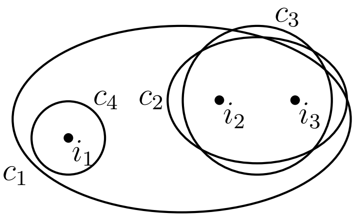
Let $N = \{i_1, \dots, i_n\}$ be a finite set of *players* or *individuals* who are able to generate some value or profit by cooperation. The cooperation takes place within groups which may overlap. For this, let a finite set $M = \{c_1, \dots, c_m\}$ of *connections* or *coalitions* be given. The elements of M are interpreted as projects or departments of an organization, for example, and the players are members of these projects.

Definition 2.1. Let $h : M \longrightarrow 2^N$ be an arbitrary mapping, where 2^N is the power set of N. The tuple (N, M, h) is a *coalitional network*.

If a coalitional network represents the structure of an organization, for instance, then the mapping h assigns to each project of the organization $c \in M$ the employees $h(c) \subseteq N$ that are working on it. In particular, from a mathematical point of view, the coalitional network (N, M, h) is simply a hypergraph (see, e.g., Berge, 1989) and the set of all coalitional networks is denoted by $\mathcal{H}$. Note that $|\mathcal{H}| = 2^{nm}$. For ease of handling, coalitional networks are also simply called networks in the following. But it should be mentioned that in literature *bilateral networks* (i.e., each connection contains exactly two players) are usually termed this way. Moreover, since N and M are not altered most of the time, a coalitional network (N, M, h) will be frequently identified with h only if no confusions can result.

Example 2.1. *As an example let $N = \{i_1, i_2, i_3\}$ and $M = \{c_1, \dots, c_4\}$, i.e., there are three players and four connections. Suppose all players are contained in c_1, the players i_2 and i_3 are moreover contained in c_2 and c_3, while c_4 only contains i_1.*
A coalitional network (N, M, h) describing this situation formally would be given by

$$h(c) = \begin{cases} \{i_1, i_2, i_3\}, & \text{if } c = c_1 \\ \{i_2, i_3\}, & \text{if } c \in \{c_2, c_3\} \\ \{i_1\}, & \text{if } c = c_4. \end{cases}$$



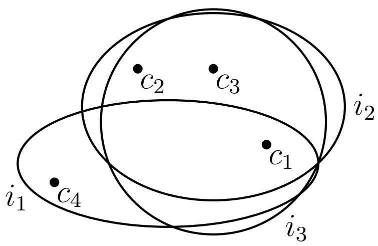
$$\begin{array}{c} \\ i_1 \\ i_2 \\ i_3 \end{array} \begin{array}{c} \begin{array}{cccc} c_1 & c_2 & c_3 & c_4 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

Figure 2.1: The coalitional network in Example 2.1

The main difference to Myerson's Communication Structures (Myerson, 1980) is that here for $c \in M$ also $|h(c)| = 1$ is allowed: that is, it is possible that a connection contains only one player. If $c \in M$ represents a project or the like and the players are employees, then it is of course possible that an employee is working on the project alone. Therefore, these structures are not excluded a priori.[1] A further important special case is $|h(c)| = 0$, i.e., it is also possible that a connection is empty or, in other words, no worker is assigned to the project. The coalitional network $h^{\varnothing}$ with $h^{\varnothing}(c) = \varnothing$ for all $c \in M$ is called *empty network*.

One of the most basic and important insights in Hypergraph Theory is the existence of a unique dual hypergraph or dual network, respectively (see, e.g., Berge, 1989).

Definition 2.2. Let $(N, M, h) \in \mathcal{H}$ be a coalitional network. The corresponding *dual network* $\big(M, N, h^*\big)$ is given by $h^*(i) = \{m \in M \mid i \in h(m)\}$.



$$\begin{array}{c} \\ c_1 \\ c_2 \\ c_3 \\ c_4 \end{array} \begin{array}{c} \begin{array}{ccc} i_1 & i_2 & i_3 \end{array} \\ \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array}$$

Figure 2.2: The dual network in Example 2.1

For given $h \in \mathcal{H}$, the dual network assigns to each player the set of connections she is contained in. As the name implies, it is a coalitional network, too, and similar to h it

[1]The main arguments for assuming that each link has to contain at least two players are often based on communication issues, like that at least two players are needed to confer, for example. However, in this case it is still possible to adjust the incentives for forming connections of size one accordingly. For instance, by assuming that these connections or departments, respectively, generate no or even negative profit (see Section 2.3 for a formal introduction of players' payoffs). For an extensive discussion of this issue see Aumann and Drèze (1974).

will be simply denoted by h^* in the following. Let $\mathcal{H}^*$ denote the set of dual networks. Note that for all $i \in N$ and $c \in M$, $c \in h^*(i)$ if and only if $i \in h(c)$. This implies $(h^*)^* = h$ and, thus, the dual network is uniquely determined. Phrased differently, the dual network is simply an alternative way of representing the network structure. However, the salient point is that from an economic perspective the players and connections interchange roles. In the dual network, the players link the connections and vice versa.

The following notions and definitions are variations or straightforward extensions from Network Theory (cf. Jackson, 2008). A network $h \in \mathcal{H}$ is a *subnetwork* of $h' \in \mathcal{H}$ (denoted by $h \subseteq h'$) if each connection in h is contained in the corresponding connection of h', i.e., $h(c) \subseteq h'(c)$ for all $c \in M$. An important special case of subnetworks is the *restriction* to a certain set of individuals. Given $S \subseteq N$, the subnetwork $h|_S$ is obtained from h by deleting in all connections all players who are not contained in S. Formally: $h|_S(c) = h(c) \cap S$ for all $c \in M|_{h,S}$. Given $h \in \mathcal{H}$, the *size* η_h of the coalitional network is the aggregated cardinality of all connections: i.e., $\eta_h := \sum_{c \in M} |h(c)|$. Two players $i, j \in N$ are called *neighbors* or *adjacent* if there exists $c \in M$ with $i, j \in h(c)$. Furthermore, the *degree* $\deg_i(h)$ of the player is the number of connections she is contained in: i.e., $\deg_i(h) = |\{c \in M \mid i \in h(c)\}|$. Note that the aggregated degree equals exactly the size of the network, that is, $\sum_{i \in N} \deg_i(h) = \eta_h$. If $i \in N$ has a degree of zero, she is said to be *isolated* and a player who is not isolated is *active*. The set of active players is $N(h) = \{i \in N \mid \deg_i(h) \geq 1\}$.
For two players $i, j \in N$ an *i-j-walk* in the coalitional network $h \in \mathcal{H}$ is a sequence of players $(i_0, \ldots, i_k)$ with $i = i_0, j = i_k$ and i_l is adjacent to i_{l+1} for all $0 \leq l < k$. A nonempty set of players $S \subseteq N$ is said to be *connected* if for any pair $i, j \in S$ there exists a walk between them. Furthermore, S is called *association* if it is maximally connected: i.e., it is connected and for all i, j with $i \in S$ and $j \notin S$ there is no i-j-walk in h. A nonempty network $\bar{h} \in \mathcal{H}$ is called *component* of $h \in \mathcal{H}$ if there exists an association $A \subseteq N$ with $\bar{h} = h|_A$. In the following, let $\mathcal{A}(h)$ be the set of all associations and $\mathcal{C}(h)$ the set of all components. Note that an association may consist of only one player who may have a degree of zero.

Definition 2.3. A *value function* is a mapping $v : \mathcal{H} \longrightarrow \mathbb{R}$ assigning a real number to each network $h \in \mathcal{H}$, where $v(h)$ is interpreted as the *worth* of the network. Thereby it is assumed that each value function is *normalized*, i.e., $v(h^{\varnothing}) = 0$.

The worth of a network is the total value of cooperation. If $h \in \mathcal{H}$ represents the structure of some organization or the like, $v(h)$ might be the organization's profit or budget. Let $\mathcal{V}$ denote the set of value functions. For each $v \in \mathcal{V}$ the corresponding dual value function $v^* : \mathcal{H}_{N,2^M} \longrightarrow \mathbb{R}$ is obtained by assigning to each dual network h^* the worth of its "primal" one: $v^*(h^*) := v\left((h^*)^*\right) = v(h)$. This induces, in particular, a bijection between $\mathcal{V}$ and the set of the dual value functions $\mathcal{V}^*$. For avoiding notational inconveniences, in the following $\mathcal{V}$ and $\mathcal{V}^*$ will be identified with each other. That is, no distinction will be made between the both sets and both of them will just be denoted by $\mathcal{V}$. Phrased differently, $v \in \mathcal{V}$ implies the value function is the primal one if v is applied to a network $h \in \mathcal{H}$ and it implies v is the dual value function if v is applied to the dual network $h^* \in \mathcal{H}$. However, because both networks receive the same worth no confusions should arise.

A specific example are *basic value functions* (e.g., Jackson, 2005). Given $h \in \mathcal{H} \setminus \{h^{\varnothing}\}$, the corresponding basic value function $1^h \in \mathcal{V}$ is defined by

$$1^h(h') = \begin{cases} 1, & \text{if } h \subseteq h' \\ 0, & \text{if } h \nsubseteq h'. \end{cases}$$

The interpretation is that a coalitional network $h' \in \mathcal{H}$ generates a positive worth if and only if it contains the specific structure h.

Let $\mathcal{CA} \subseteq \mathcal{V}$ denote the special class of component-additive value functions. A value function $v \in \mathcal{CA}$ is *component-additive* if $\sum_{\bar{h} \in \mathcal{C}(h)} v\left(\bar{h}\right) = v(h)$ for all coalitional networks $h \in \mathcal{H}$. Here, the total productivity of a coalitional network is simply the aggregated worth of all its components (cf. Jackson and Wolinsky, 1996). The productivity of a component is therefore not exogenously influenced by the other components.

Remark 2.1. Traditionally, a TU game is a pair (N, γ) consisting of the set of players N and a *characteristic function* γ assigning a value to each *coalition* $S \subseteq N$, i.e., $\gamma : 2^N \longrightarrow \mathbb{R}$ (see, e.g., Roth, 1988). On the other hand, a (bilateral) network game is a pair (N, w) where w assigns a value to each *bilateral network*, i.e., $w : 2^G \longrightarrow \mathbb{R}$ with $G = \{S \subseteq N \,|\, |S| = 2\}$ (see, e.g., Jackson, 2008). The main idea for combining both approaches within the extended framework considered here is to exploit the fact that both domains, coalitions and bilateral networks, can be embedded into the set of coalitional networks. For instance, if $|M| = 1$, i.e., if there is only one connection, this connection can be interpreted as a coalition and value functions can be identified with characteristic functions. Thus, the model is indeed an extension of TU games. Moreover, a network game (N, w) can also be represented in a canonical way: let $M = G$ and for each $g \in 2^G$ define

$$h^g(\{i,j\}) = \begin{cases} \{i,j\}, & \text{if } \{i,j\} \in g \\ \varnothing, & \text{if } \{i,j\} \notin g. \end{cases}$$

That is, i and j are adjacent in h^g if and only if they are linked in g. Note that the value or worth, respectively, of each bilateral network only depends on whether certain players are linked but not on in which way they are connected. Therefore, if $v^w \in \mathcal{V}$ satisfies not only $v^w(h^g) = w(g)$ for all $g \in 2^G$ but also $v^w(h^g \circ \pi) = v^w(h^g)$ for each permutation $\pi : M \longrightarrow M$, this value function de facto represents w.[2]

Remark 2.2. Basic value functions are a straightforward extension of unanimity games (cf. Shapley, 1953). It is well-known from literature that these games form a basis of the space of characteristic functions. Analogously it can be shown that the set of basic value functions forms a basis of $\mathcal{V}$. More precisely, each $v \in \mathcal{V}$ can be written as

$$v = \sum_{h \in \mathcal{H} \setminus \{h^{\varnothing}\}} \Delta^h(v) 1^h, \quad \text{where the} \quad \Delta^h(v) = \sum_{\bar{h} \subseteq h} (-1)^{\eta_h - \eta_{\bar{h}}} v(\bar{h})$$

are the Harsanyi coefficients (cf. Harsanyi, 1959). This representation of value functions will be helpful in the proofs.

Remark 2.3. Note that $\mathcal{CA} = \{v \in \mathcal{V} \mid v \text{ is component-additive}\}$ is a subspace of $\mathcal{V}$ and $1^h \in \mathcal{CA}$ if and only if $|\mathcal{C}(h)| = 1$ (cf. van den Nouweland and Slikker, 2012). Moreover, since the operator $* : \mathcal{V} \longrightarrow \mathcal{V}$ with $v \longmapsto v^*$ is not only bijective but also linear and $|\mathcal{C}(h)| = 1$ is equivalent to $|\mathcal{C}(h^*)| = 1$, $v \in \mathcal{CA}$ is component-additive if and only if $v^* \in \mathcal{CA}$ has this feature, too.

2.3 Allocation Rules

Given the productivity of cooperation, the next step is to allocate the worth among the players. Therefore, the pair (N, M) is an *allocation problem* (note that $\mathcal{V} = \mathcal{V}(N, M)$ is uniquely determined by N and M). Most of the solutions analyzed in Cooperative Game Theory or Network Theory distribute the worth in a player-based way, i.e., the payoff is associated to players directly.

[2] As usual, $\circ$ denotes the composition of the mappings h^g and π.

Definition 2.4. A *player-based allocation rule* X assigns to each allocation problem (N, M) a tuple of mappings $(X_i)_{i\in N}$ with $X_i : \mathcal{V} \longrightarrow \mathcal{V}$ and $\sum_{i\in N} X_i(v) = v$ for all $v \in \mathcal{V}$.[3]

A player-based allocation rule X (or $(X_i)_{i\in N}$, respectively) decomposes each value function $v \in \mathcal{V}$ into individual payoff functions $v_i^X := X_i(v) \in \mathcal{V}$. Phrased differently, given the total productivity $v(h)$ of a network h, each player receives her share of the payoff $v_i^X(h) = (X_i(v))(h)$. The central assumption in Definition 2.4 is that allocation rules are *balanced*, i.e., $\sum_{i\in N} v_i^X(h) = v(h)$ for all $h \in \mathcal{H}$ and $v \in \mathcal{V}$. Thus, the value of cooperation is always distributed completely among the players.[4] The sets of all player-based allocation rules will be denoted by $\mathcal{X}$. Note that each allocation rule $Z \in \mathcal{X}$ assigns payoff functions $(v_c^Z)_{c\in M}$ to the dual problem (M, N), too. This is due to the fact that in the dual network the connections are considered as players and vice versa.

Example 2.2. *A possible way to allocate the worth of a network is to distribute it equally among all active players:*

$$v_i^{ED}(h) = (ED_i(v))(h) := \begin{cases} \frac{v(h)}{|N(h)|}, & \textit{if } i \in N(h) \\ 0, & \textit{if } i \notin N(h), \end{cases}$$

where $h \in H_{M,2^N}$ *and* $v \in \mathcal{V}$. *For instance, in Example 2.1 this yields* $v_i^{ED}(h) = \frac{v(h)}{4}$ *for all* $i \in N$. *Analogously, in the corresponding dual network* h^* *each connection* $c \in M$ *receives* $v_c^{ED}(h^*) = \frac{v(h^*)}{3} = \frac{v(h)}{3}$. *The allocation rule* $ED \in \mathcal{X}$ *will be called* Equal Division Solution *in the following.*

Distributing the worth in two stages cannot be modeled explicitly by means of player-based allocation rules since, as mentioned before, in this case it is first distributed

[3] In their seminal contribution, Jackson and Wolinsky (1996) introduce (player-based) allocation rules in a slightly different way. They define an allocation rule A to be a mapping $A : \mathcal{H} \times \mathcal{V} \longrightarrow \mathbb{R}^n$ with $\sum_{i\in N} A_i(h, v) = v(h)$ for all $h \in \mathcal{H}$ and $v \in \mathcal{V}$. The interpretation is, of course, the same. Given a network $h \in \mathcal{H}$ and a value function $v \in \mathcal{V}$, $A_i(h, v)$ is agent i's share of the network's value $v(h)$. Using the formulation introduced here is just for technical reasons and will be helpful in the following.

[4] In contrast to efficiency from Cooperative Game Theory (an allocation rule is said to be efficient if it always distributes the value that could be reached by the grand coalition N (see, e.g., van den Brink, 2007)), it is not implicitly assumed that the grand coalition or the complete network, respectively, generically generates the highest value. Therefore, the assumption of balancedness is less restrictive and more natural.

among the connections and subsequently within each connection. Thus, it is not associated to the players directly but, if the connections represent projects, for example, each player receives her share of each project's budget. Capturing this aspect requires a refined concept of allocation rules.

Definition 2.5. A *position* is a pair $ic := (i,c) \in N \times M$. A *position-based allocation rule* Y assigns a family of functions $Y = (Y_{ic})_{ic \in N \times M}$ to each allocation problem (N,M) with $Y_{ic} : \mathcal{V} \longrightarrow \mathcal{V}$ and $\sum_{c \in M, i \in N} Y_{ic}(v) = v$ for all $v \in \mathcal{V}$. Moreover, given the dual problem (M,N), each Y is assumed to satisfy $(Y_{ci}(v^*))^* = Y_{ic}(v)$ for all $v \in V$.

The interpretation is basically the same as the interpretation of player-based allocation rules. But in addition there is a further symmetry axiom imposed: it requires that each position in the dual problem receives the same payoff as in the primal one. The set of position-based allocation rules is $\mathcal{Y}$. The salient point is that every position-based allocation rule $Y \in \mathcal{Y}$ induces two player-based allocation rules $X^Y \in \mathcal{X}$ and $Z^Y \in \mathcal{X}$ in a straightforward way. Given $v \in \mathcal{V}$, each player/connection receives the aggregated payoff of all her/its positions:

1. $(X_i^Y)_{i \in N}$ is given by $v_i^Y := X_i^Y(v) = \sum_{c \in M} Y_{ic}(v)$ and
2. $(Z_c^Y)_{c \in M}$ by $v_c^Y := Z_c^Y(v) = \sum_{i \in N} Y_{ic}(v)$.

2.3.1 A Specific Two-stage Procedure

Position-based allocation rules allow for more flexibility in tracking how the worth of a network is distributed among the players. In particular, they allow modeling two-stage allocation procedures explicitly. To fix ideas, let a pair of player-based allocation rules $(X,Z) \in \mathcal{X} \times \mathcal{X}$ be given and for each allocation problem (N,M) define $Y \in \mathcal{Y}$ via $Y_{ic} := X_i \circ Z_c$ for all $ic \in N \times M$. The interpretation is as follows:

1. The allocation rule Z specifies the payoff at the first stage. Given a network $h \in \mathcal{H}$, the connections play a dual game in h^* and they are paid according to Z. That is, connection $c \in M$ receives $v_c^Z(h^*) = (Z_c(v))(h^*)$.
2. Then, in a second step, for each $c \in M$ the value $v_c^Z(h^*)$ is distributed among the players according to $X \in \mathcal{X}$. More specifically, player i's share of connection c is $v_{ic}^Y(h) = \big(X_i\big(v_c^Z\big)\big)(h) = (X_i(Z_c(v)))(h) = ((X_i \circ Z_c)(v))(h)$.

Consequently, a position-based allocation rule $Y \in \mathcal{Y}_{M,2^N}$ is said to be *induced by two player-based allocation rules* if there exists a pair $(X,Z) \in \mathcal{X} \times \mathcal{X}$ with $Y_{ic} := X_i \circ Z_c$ for each allocation problem (N,M) and all $ic \in N \times M$. Let $\mathcal{Y}^{2ply} := \{Y \in \mathcal{Y} \,|\, \mathcal{Y}$ is induced by two player-based allocation rules$\}$. As will be shown later, this is just a special case of a broader class of two-stage allocation procedures. Nevertheless, it possesses striking and remarkable characteristics which are worth studying in more detail.

Lemma 2.1. Let an allocation problem (N,M) be given. If a position-based allocation rule $Y \in \mathcal{Y}^{2ply}$ is induced by two player-based allocation rules $X \in \mathcal{X}$ and $Z \in \mathcal{X}$, then $Z_c = Z_c^Y$ for all $c \in M$.

Proof. This is a straightforward implication of the balancedness assumption. In particular, in the second step the players distribute exactly the first-stage value among each other:

$$v_c^Y = \sum_{i \in N} Y_{ic}(v) = \sum_{i \in N} (X_i(Z_c(v))) = Z_c(v) = v_c^Z \quad \text{for all } i \in N \text{ and } v \in \mathcal{V}.$$

□

Thus, if $Y \in \mathcal{Y}^{2ply}$ is induced by two player-based allocation rules, the first-stage allocation Z and the connections' values determined by Z^Y do not differ. Generically, this is not necessarily true for X and X^Y. Redistributing the value at the first stage might have a significant impact on the payoff of the players. However, there is an important class of allocation rules where both ways of allocating the worth coincide.

Definition 2.6. A player-based allocation rule $X \in \mathcal{X}$ is *additive* (ADD) if $X(v + v') = X(v) + X(v')$ for all (N,M) and $v, v' \in \mathcal{V}$. Additivity for position-based allocation rules is defined analogously.

Additive allocation rules exhibit a non-externalities property. The payoff generated by a value function $v \in \mathcal{V}$ is not influenced by other value functions.

Proposition 2.1. Let an allocation problem (N,M) be given and suppose $Y \in \mathcal{Y}^{2ply}$ is induced by two player-based allocation rules $X \in \mathcal{X}$ and $Z \in \mathcal{X}$. If X is additive, $X_i = X_i^Y$ for all $i \in N$.

Proof. If X is additive, then

$$v_i^Y = \sum_{c \in M} Y_{ic}(v) = \sum_{c \in M} (X_i(Z_c(v))) = X_i \circ \left(\sum_{c \in M} Z_c(v) \right) = X_i(v) = v_i^X$$

for all $i \in N$ and $v \in \mathcal{V}$. □

Proposition 2.1 shows that the final outcome obtained by integrating an additive allocation rule $X \in \mathcal{X}$ into a two-stage allocation procedure as introduced above does not differ from applying X in a player-based way. Moreover, this is completely independent of the first-stage allocation rule $Z \in \mathcal{X}$. It does not matter in which way the value is redistributed at the first stage, the final outcome is always the same. In particular, these considerations allow some interesting implications: Suppose a coalitional network represents a complexly structured firm or organization and it is more convenient to pay the players for each project separately instead of calculating the payoff of all players as a function of the organization's revenue. If an additive allocation rule is chosen, the players' payoff is not changed. This motivates the following definition:

Definition 2.7. A player-based allocation rule $X \in \mathcal{X}$ is *invariant under redistribution* if for all allocation problems (N, M) and $v \in \mathcal{V}$,

$$X_i(v) = \sum_{c \in M} X_i\left(Z_c(v)\right) \quad \text{for all } Z \in \mathcal{X}. \tag{2.1}$$

As it has been shown in Proposition 2.1, additive allocation rules satisfy this criterion. It turns out that these are indeed the only ones. The proof of the following result can be found in the appendix.

Proposition 2.2. A player-based allocation rule $X \in \mathcal{X}$ is invariant under redistribution if and only if it is additive.

Although Proposition 2.2 might be slightly surprising at first sight, again it is just a direct implication of the balancedness assumption. This together with additivity immediately implies Equation (2.1). However, even if Proposition 2.2 is not very complicated from a mathematical point of view, in the context of coalitional networks it gives rise to an alternative interpretation of additivity. It allows for arbitrary redistributions of the worth without changing the final outcome of the players. In fact, many allocation rules discussed in the literature are additive, such as the Shapley Value, the Equal Division Solution or the Position Value, to name but a few.

Before completing this subsection there remains a further question which requires special attention. So far the discussion concentrated on properties of allocation rules in $\mathcal{Y}^{2ply}$ without explicitly characterizing this particular class. In other words: Which position-based allocation rules are actually induced by two player-based allocation rules? In order to approach this question, let $Y \in \mathcal{Y}$ and suppose there exist $X \in \mathcal{X}$

and $Z \in \mathcal{X}$ with $Y_{ic} = X_i \circ Z_c$ for all allocation problems (N, M). This, in particular, implies that each position's value only depends on the corresponding connection's value generated at the first-stage but not on the connection itself. That is, if the first-stage outcomes v_c^X and $\bar{v}_{\bar{c}}^X$ of two connections $c, \bar{c} \in M$ for two allocation rules $v, \bar{v} \in \mathcal{V}$ coincide, the payoff player i receives for both positions has to be the same as well, i.e, $X_i(v_c^X) = X_i(\bar{v}_{\bar{c}}^X)$. By applying Lemma 2.1 this independence property can be restated as follows: If the values of two connections coincide, each player receives the same payoff from both of them. In fact, this property is not only necessary but also sufficient.

Proposition 2.3. A position-based allocation rule $Y \in \mathcal{Y}$ is induced by two player-based allocation rules if and only if it satisfies the following independence property: Given an allocation problem (N, M), if $v_c^Y = \bar{v}_{\bar{c}}^Y$ for $c, \bar{c} \in M$ and $v, \bar{v} \in \mathcal{V}$, then $v_{ic}^Y = \bar{v}_{i\bar{c}}^Y$ for all $i \in N$.

The proof can be found in the appendix. A consequence of this independence property is that players might receive payoff from connections they are not a member of. Consider for example the Equal Division Solution $ED \in \mathcal{X}$ which is additive and let an allocation problem (N, M) and a network $h \in \mathcal{H}$ be given. If $Z \in \mathcal{X}$ is a first-stage allocation rule, generically it is possible that $((ED_i \circ Z_c)(v))(h) \neq 0$ even if $i \notin h(c)$. This is due to the fact that v_c^Z depends on the whole network and not only on c. Of course, in certain situations it is quite reasonable that players receive payoff from connections they are not a member of, like in the case of externalities, for example. However, in many environments it seems to be more convincing that the value of a connection is distributed only among the corresponding members. This aspect is not captured so far and it requires a higher degree of flexibility. That is, instead of applying the same second-stage allocation rule $X \in \mathcal{X}$ for all $c \in M$ it is necessary to use for each connection a specifically designed $X^c \in \mathcal{X}$.

2.3.2 Reduced Allocation Problems

Consider a similar situation as before. Again, the connections first receive a certain payoff, but now the players are supposed to face reduced problems of allocating the value within each connection. The main idea is that the corresponding members take the rest of the network as given and compare the prospects if only their connection changes. To define this formally, assume that there is given a coalitional network $h \in \mathcal{H}$, a value function $v \in \mathcal{V}$, and a first-stage allocation rule $Z \in \mathcal{X}$. For each set of players $S \subseteq N$ the network $h|_{c,S} \in \mathcal{H}$ is obtained from h by deleting all players in $h(c)$ that are not contained in S while the other links remain the same, i.e.,

$$h|_{c,S}(c') = \begin{cases} h(c'), & \text{if } c' \neq c \\ h(c) \cap S, & \text{if } c' = c \end{cases}$$

for all $c, c' \in M$. Furthermore, let $h|^{*}_{c,S}$ denote the corresponding dual network. This substructure allows designing a reduced allocation problem $(N, \{c\})$ within connection $c \in M$. Each subset of players $S \subseteq N$ can be interpreted as a coalitional network with only one connection (cf. Remark 2.1), and the corresponding value function $w^{c,h}_{v^Z} \in \mathcal{V}_{\{c\},2^N}$ of the reduced problem is given by $w^{c,h}_{v^Z}(S) := v^Z_c\left(h|^{*}_{c,S}\right)$. The players only take into account in which way the first-stage payoff changes if deviations within c occur while the other connections remain the same. In order to guarantee that $w^{c,h}_{v^Z} \in \mathcal{V}_{\{c\},2^N}$ is well-defined, that is, to guarantee that $w^{c,h}_{v^Z}$ is normalized, in the following it will always be required that $Z \in \mathcal{X}$ satisfies the active players axiom.

Definition 2.8. A player-based allocation rule $X \in \mathcal{X}$ satisfies the *inactive players axiom* (IPLY) if for each allocation problem (N, M), $i \notin N(h)$ implies $v^X_i(h) = 0$ for all $h \in \mathcal{H}$ and $v \in \mathcal{V}$.

If a player-based allocation rule satisfies this axiom, players who are not active get zero payoff. In the dual game considered here, this means that empty connections receive a worth of zero.

Now suppose $X^{re} \in \mathcal{X}$ is an allocation rule applied to the reduced problem $(N, \{c\})$. This induces finally a second-stage allocation rule $X^c \in \mathcal{X}$ via $\big(X^c_i\big(v^Z_c\big)\big)(h) := \big(X^{re}_i\big(w^{c,h}_{v^Z}\big)\big)(h(c))$ for all $i \in N$.[5] Note that although the main motivation was to construct a position-based allocation rule which gives zero payoff to positions not contained in the network, up to here it is not required that this condition is indeed satisfied. However, assuming additionally for all connections $c \in M$ that the allocation rule applied to $(N, \{c\})$ satisfies the inactive players axiom is sufficient to achieve the desired effect.

Of course, the main idea of considering reduced problems of allocating the value is similar to Subsection 2.3.1, but the main difference is that now the allocation rules in the second step do not necessarily have to be the same for all connections. Therefore, the corresponding two-stage procedure has the following form:

[5]For further allocation problems $(N', M') \neq (N, M)$ it would be possible to proceed analogously.

1. Given a value function $v \in \mathcal{V}$, the value of $c \in M$ is determined according to a first-stage allocation rule $Z \in \mathcal{X}$.

2. Then v_c^Z is allocated according to $X^c \in \mathcal{X}$ for each $c \in M$, i.e., player i receives $X_i^c\big(v_c^Z\big)$.

In the following, a position-based allocation rule $Y \in \mathcal{Y}$ is said to be *implemented in two stages* if for each allocation problem (N, M) there exist $X^c \in \mathcal{X}$ for each $c \in M$ and $Z \in \mathcal{X}$ such that $Y_{ic} = X_i^c \circ Z_c$ for all $i \in N$ and $c \in M$. Let $\mathcal{Y}^{2stg} \subseteq \mathcal{Y}$ be the set of position-based allocation rules satisfying this criterion, i.e., $\mathcal{Y}^{2stg} = \{Y \in \mathcal{Y} \mid Y \text{ is implemented in two stages}\}$. Note that $\mathcal{Y}^{2ply} \subseteq \mathcal{Y}^{2stg}$.

Lemma 2.2. Let an allocation problem (N, M) be given and assume a position-based allocation rule $Y \in \mathcal{Y}^{2stg}$ is implemented in two stages, where $Z \in \mathcal{X}$ is the first-stage allocation rule. Then $Z_c = Z_c^Y$ for all $c \in M$.

Proof. The reasoning is completely analog to Lemma 2.1. □

In particular, it does not matter whether for each connection a specifically designed allocation rule is used or whether it is always the same one. The aggregated value is the same in any case due to the balancedness assumption. Moreover, although $\mathcal{Y}^{2stg}$ is a broader class of allocation rules than $\mathcal{Y}^{2ply}$, it is possible to characterize it by a similar independence property.

Proposition 2.4. A position-based allocation rule $Y \in \mathcal{Y}$ is implemented in two stages if and only if it satisfies the following independence property: Given an allocation problem (N, M), if $v_c^Y = \bar{v}_c^Y$ for $c \in M$ and $v, \bar{v} \in \mathcal{V}$, then $v_{ic}^Y = \bar{v}_{ic}^Y$ for all $i \in N$.

The proof proceeds analogously to the one of Proposition 2.3 and the interpretation is similar as well. The difference is that here, the payoff of each position has to be independent only of how the connection's value is generated but not of the connection itself. This is, of course, due to the fact that the second-stage allocation rule now may depend on c.
The main relevance of the previous proposition is that it characterizes all two-stage allocation procedures in a natural way. Therefore, implicitly it also provides a tool for checking whether position-based allocation rules belong to this class or not. The remainder of this paper is devoted to the study of particular examples, like the Position Value, for instance. It also includes a general discussion of properties of allocation rules.

2.4 Some Particular Allocation Rules

Before analyzing some specific two-stage allocation procedures, first the focus lies on player-based allocation rules. The economic implications of this intermediate step are actually just straight-forward extensions from Network Theory and Cooperative Game Theory. Nevertheless, they strongly support the intuition of the two-stage allocation procedures analyzed later in Subsection 2.4.2 and provide a technical substructure for them.

2.4.1 Player-based Allocation Rules

One specific player-based allocation rule, the Equal Division Solution, has already been introduced in Example 2.2. Another well-known example from literature is the Myerson Value which has been established by Myerson (1977) and extended to bilateral networks by Jackson and Wolinsky (1996). The main motivation of this allocation rule is to pay each player her expected marginal contribution to a (bilateral) network. Clearly, this idea can be adopted completely analogously to the setting considered here. Formally this means that given an allocation problem (N, M), for each $h \in \mathcal{H}$ and $v \in \mathcal{V}$ the *Myerson Value* (for coalitional networks) $MV\mathcal{X}$ is defined by

$$v_i^{MV}(h) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - (|S| + 1))!}{n!} \cdot (v(h|_{S+i}) - v(h|_S)) \quad \text{for all } i \in N.^6$$

At first sight the Equal Division Solution and the Myerson Value seem to be quite different and a priory it is not clear which one is appropriate for which situation. To gain a better understanding, it is necessary to analyze which properties both allocation rules have and in which way they differ.

Definition 2.9. Let an arbitrary allocation problem (N, M) be given. A player-based allocation rule $X \in \mathcal{X}$ satisfies ...

(CBAL) ... *component-balancedness* if $\sum_{i \in A} v_i^X(h) = v(h|_A)$ for each network $h \in \mathcal{H}$, every component-additive value function $v \in \mathcal{CA}$, and all associations $A \in \mathcal{A}(h)$.

(SYM) ... *symmetry of active players* if for $i, j \in N(h)$, $v(h|_{S+i}) = v(h|_{S+j})$ for all $S \subseteq N(h) \setminus \{i, j\}$ always implies $v_i^X(h) = v_j^X(h)$.

[6] The Myerson Value is, roughly speaking, an extension of the *Shapley Value* (Shapley, 1953). In fact, if Φ denotes the Shapley Value, $v_i^{MV}(h) = \Phi_i\left(w_v^h\right)$ where $w_v^h(S) := v(h|_S)$ for all $S \subseteq N$.

(NULI) ...the *nullifying player property* if $v(h|_{S+i}) = 0$ for all $S \subseteq N(h) \setminus \{i\}$ implies $v_i^X(h) = 0$.

(NULL) ...the *null player property* if $v(h|_{S+i}) - v(h|_S) = 0$ for all $S \subseteq N(h)$ implies $v_i^H(h) = 0$.

The previous notions are standard in the literature of Cooperative Game Theory and Network Theory and therefore the interpretations do not change. Given a *component-additive* value function, an allocation rule is component-balanced if it distributes the value of each component only among its members. *Symmetry of active players* requires that two players receive the same payoff whenever their contributions to the network are equal. An allocation rule satisfies the *nullifying player property* if a player receives no payoff whenever her presence in a network induces a value of zero. Deegan and Packel (1978) call these individuals zero players. In contrast to this, the *null player property*, also known as the dummy axiom, focuses on the marginal contribution of a player to the network. The usual interpretation is that a player has no bargaining power if her marginal contribution to the network is always zero and, thus, she should receive no payoff. Note that this implies the inactive players axiom.

In the context of Cooperative Game Theory, van den Brink (2007) proved that the Shapley Value and the Equal Division Solution satisfy similar axioms. In fact, it is possible to extend this insight to the more general setting considered here.

Proposition 2.5. $X \in \mathcal{X}$ satisfies (ADD), (SYM), (IPLY), and (NULI) if and only if $Y = ED$.

For the proof (and also for the proofs of the following results) refer to the appendix. In principle it proceeds in a standard fashion. More precisely, given an allocation problem (N, M), due to additivity it is sufficient to focus on basic value functions which form a basis of $\mathcal{V}$ (recall Remark 2.2).

Proposition 2.6. $X \in \mathcal{X}$ satisfies (ADD), (SYM), and (NULL) if and only if $Y = MV$.

Since the Myerson Value is an extension of the Shapley Value (Shapley, 1953), it can be characterized by similar axioms. Analogously to the characterization of the Equal Division Solution, the proof of Proposition 2.6 also exploits additivity and concentrates on basic value functions.

The previous results show that if an allocation rule is supposed to satisfy additivity, symmetry of active players, and the inactive players axiom, the final outcome strongly depends on the decision whether more importance is attached to nullifying or to null players. In one case the Myerson Value and in the other the Equal Division Solution is obtained.

In fact, in Cooperative Game Theory the null player property is often used to characterize solutions. Besides the Myerson and the Shapley Value, also Owen's Value (Owen, 1977), for example, satisfies this axiom. It allows some noteworthy implications which should not be neglected here:

Lemma 2.3. If $X \in \mathcal{X}$ satisfies (ADD) and (NULL), it also satisfies (CBAL).

Proposition 2.6 and Lemma 2.3 immediately imply the following result.

Proposition 2.7. The Myerson Value is component-balanced.

The intuition of Lemma 2.3 is straightforward. If a value function is component-additive, the value of a component does not depend on the others. Therefore, the marginal contribution of a player to other components is always zero. The result then follows from the null player property and because the individual payoff functions can be decomposed appropriately by additivity.[7]

2.4.2 Position-based Allocation Rules

All the properties introduced in the previous subsection focus on the players directly but not on the positions. The marginal contribution, for instance, is determined only for subnetworks where player $i \in N$ completely drops out of the network. In this case, she totally stops cooperating. In economic life, however, if the connections represent projects or the like, it might happen that a player only leaves some of her projects but not all of them. Translated to the model this would mean that she stops cooperating only partially. To fix ideas, consider the position-based variations of the properties introduced in the previous sections:

Definition 2.10. Let an arbitrary allocation problem (N, M) be given. A position-based allocation rule $Y \in \mathcal{Y}$ satisfies ...

[7]For bilateral networks Jackson and Wolinsky (1996), who refer to Myerson (1977), have already established Proposition 2.7. Although they have shown it in a different context, and thus in a different way, some elements of the proof of Lemma 2.3 are adopted from these works.

(CBAL') ... *component-balancedness* if $\sum_{i\in A}\sum_{c:h(c)\subseteq A} v_{ic}^{Y}(h) = v(h|_A)$ for all $h \in \mathcal{H}$, every component-additive value function $v \in \mathcal{CA}$, and all associations $A \in \mathcal{A}(h)$.

(IPOS) ... the *inactive positions axiom* if $i \notin h(c)$ implies $v_{ic}^{Y}(h) = 0$ for all $v \in \mathcal{V}$.

(cwSYM) ... *symmetry of active players connection-wise* if for each $c \in M$ and $i,j \in h(c)$ from $v_c^{Y}\left(h|_{c,S\cup\{i\}}^{*}\right) = v_c^{Y}\left(h|_{c,S\cup\{j\}}^{*}\right)$ for all $S \subseteq h(c) \setminus \{i,j\}$ always $v_{ic}^{Y}(h) = v_{jc}^{Y}(h)$ follows.

(cwNULI) ... *connection-wise* the *nullifying player property* if for each $c \in M$ from $v_c^{Y}\left(h|_{c,S}^{*}\right) = 0$ for all $S \subseteq N(h) \setminus \{i\}$ always $v_{ic}^{Y}(h) = 0$ follows.

(cwNULL) ... *connection-wise* the *null player property* if for each $c \in M$ from $v_c^{Y}\left(h|_{c,S\cup\{i\}}^{*}\right) = v_c^{Y}\left(h|_{c,S}^{*}\right)$ for all $S \subseteq h(c) \setminus \{i\}$ always $v_{ic}^{Y}(h) = 0$ follows.

Component-balancedness and the *inactive positions axiom* are analogously defined as the axioms for player-based allocation rules and therefore the interpretations are the same as well. However, here the focus lies on the positions but not on the players. For example, *(cwNULL)* concentrates on the marginal contributions of the players to single connections. Given this axiom, player $i \in N$ receives nothing from $c \in M$ if her marginal contribution to the connection is zero. Similar considerations also apply if the position-based allocation rule satisfies *connection-wise the nullifying player property* or *symmetry of active players*.

One of the most prominent two-stage allocation rules in Network Theory is the often-studied Position Value (e.g., Borm *et al.*, 1992). The main idea is to determine the connections' values according to the Myerson Value and then, in the second step, the members of each connection apply the Equal Division Solution in order to share the corresponding value equally. Of course, this motivation can be extended to the more general setting considered here. Given an allocation problem (N, M), the *Position Value* (for coalitional networks) $PV \in \mathcal{Y}$ is given explicitly by

$$v_{ic}^{PV}(h) = \left(PV_{ic}(v)\right)(h) := \begin{cases} \frac{1}{|h(c)|} \cdot v_c^{MV}\left(h^*\right), & \text{if } i \in h(c) \\ 0, & \text{if } i \notin h(c) \end{cases}$$

for all $i \in N$ and $c \in M$. Note that each player i receives

$$v_i^{PV}(h) = \sum_{c:i\in h(c)} \frac{1}{|h(c)|} \cdot v_c^{MV}\left(h^*\right)$$

and the value of each connection is indeed given by the Myerson Value: $v_c^{PV} = v_c^{MV}$ for all $c \in M$. Taking into account the characterizations established in Propositions 2.5 and 2.6, these considerations imply that the Position Value needs to satisfy (ADD), (SYM), together with (NULL) at the first stage and (ADD), (IPOS), (cwSYM), and (cwNULI) with respect to each connection. In fact:

Proposition 2.8. Suppose $Y \in \mathcal{Y}$ satisfies (ADD), (IPOS), (cwSYM), (cwNULI), and assume $Z^Y \in \mathcal{X}$ satisfies (NULL) and (SYM). Then $Y \in \mathcal{Y}^{2stg}$ and, furthermore, $Y = PV$.[8]

Again, the proof can be found in the appendix. Showing $Y \in \mathcal{Y}^{2stg}$ is based on exploiting (cwNULI) and (ADD). By applying the latter axiom and Proposition 2.4, it is sufficient to show that whenever a connection receives a value of zero in all networks, the positions within this connection receive a payoff of zero, too. This is given, obviously, by (cwNULI). Moreover, since the first-stage allocation rule is the Myerson Value, component-balancedness also carries over to the Position Value.

Lemma 2.4. If $Y \in \mathcal{Y}$ is component-balanced both induced player-based allocation rules $X^Y \in \mathcal{X}$ and $Z^Y \in \mathcal{X}$ have this property, too.

Proof. Let (N, M) be an allocation problem and $v \in \mathcal{CA}$ be component-additive. Furthermore, let $h \in \mathcal{H}$ and $A \in \mathcal{A}(h)$. Then:

$$\sum_{i \in A} v_i^Y(h) = \sum_{i \in A} \sum_{c \in M} \underbrace{v_{ic}^Y(h)}_{=0, \text{ if } i \notin h(c)} = \sum_{i \in A,\, c: h(c) \subseteq A} v_{ic}^Y(h) = v\left(h|_A\right).$$

Analogously for $Z_c^Y = \sum_{i \in N} Y_{ci}$. □

Lemma 2.5. Let $Y \in \mathcal{Y}$ satisfy (IPOS). Then $Z^Y \in \mathcal{X}$ is component-balanced if and only if $X^Y \in \mathcal{X}$ is component-balanced.

Proof. For the proof just note that if Y satisfies the inactive positions axiom, the proof of Lemma 2.4 works in both directions. □

[8] Borm *et al.* (1992) introduce a characterization of the Position Value for a certain class of Myerson's communication situations (cf. Myerson, 1977). Interestingly, they also interpret their "arc game" as a "dual game": "[The value] of an arc is measured by means of the Shapley Value of a kind of 'dual' game [...]" (Borm *et al.*, 1992, p. 306). Later, van den Nouweland and Slikker (2012) extended this characterization to bilateral networks. However, it is not possible to extend it further to coalitional networks in a straightforward way, because they use the "superfluous link property", which is not as restrictive here as in the bilateral setting.

Proposition 2.9. Let $Y \in \mathcal{Y}$ satisfy (IPOS) and (ADD). If, furthermore, $Z^Y \in \mathcal{X}$ satisfies (NULL), then $X^Y \in \mathcal{X}$ is component-balanced.

Proof. This is a direct implication of Lemmas 2.3 and 2.5. □

Of course, Proposition 2.9 is also satisfied if X^Y and Z^Y interchange roles.

Proposition 2.10. The Position Value is component-balanced.

For bilateral networks, van den Nouweland and Slikker (2012) already established Proposition 2.10. However, to show it, the authors used "component decomposability" which is not needed here. In the context of coalitional networks, the result is a direct implication of the characterization of the Position Value.

Although the Position Value possesses interesting and desirable features, in some situations it might be inadequate. For example, using the Myerson Value first and the Equal Division Solution in the second step violates consistency. Following Hart and Kurz (1983), a two-stage allocation rule is consistent if the way of distributing the value within each connection is the same as the one used for determining the values in the dual game. Of course, there are several ways to construct consistent allocation rules, but in light of the preceding discussion, a canonical candidate is the iterated application of the Myerson Value. Therefore, consider again a two-stage procedure where the first-stage allocation rule is $MV \in \mathcal{X}$. Given an allocation problem (N, M), for each $h \in \mathcal{H}$ and $v \in \mathcal{V}$, the value function of the reduced game within $c \in M$ is then given by $w^{c,h}_{v^{MV}}(S) = v^{MV}_c\left(h|^*_{c,S}\right)$ for all $S \subseteq N$. In contrast to the Position Value, suppose now that the second-stage allocation rules are chosen in a consistent way: that is, the solutions of the reduced games are also determined according to the Myerson Value $MV \in \mathcal{X}$ for all $c \in M$. This iterated application of the Myerson Value will be denoted by $IM \in \mathcal{Y}$ and it is given by $v^{IM}_{ic}(h) := \big(MV_i\big(w^{c,h}_{v^{MV}}\big)\big)(h(c))$.
As already discussed in Section 2.3, the characterizations of the Equal Division Solution and the Myerson Value differ by only one axiom. Consequently, the same is also true for the Position Value and the iterated application of the Myerson Value. In fact, instead of the nullifying player property, IM satisfies connection-wise the null player property.

Proposition 2.11. Let $Y \in Y_{M,2^N}$ satisfy (ADD), (cwSYM), (cwNULL), and assume $Z^Y \in \mathcal{X}$ satisfies (SYM). Then $Y \in \mathcal{Y}^{2stg}$, Z^Y satisfies (NULL), and, furthermore, $Y = IM$.

Moreover, since in the first step the worth is distributed according to the Myerson Value, Lemma 2.5 can be applied again.

Proposition 2.12. IM is component-balanced.

Applying the Myerson Value iteratively is in spirit with a prominent solution concept from Cooperative Game Theory, namely with Owen's Value (Owen, 1977). In his seminal contribution, the author focused on cooperative games with a given coalition structure, that is, the players are partitioned exogenously into coalitions which may not overlap. Given this setting, Owen applied the Shapley Value iteratively, first among the coalitions and then within each coalition. Loosely speaking, this approach does not take the actual form of cooperation $h \in \mathcal{H}$ into account but instead it focuses only on the induced associations $\mathcal{A}(h)$. To fix ideas, let (N, M) be an allocation problem and let Ψ denote Owen's Value. Moreover, for given $v \in \mathcal{V}$, let $w_v^h(S) := v(h|_S)$ be the worth of $h \in \mathcal{H}$ restricted to $S \subseteq N$. Then Owen's Value can be extended analogously to the Myerson Value via $v_i^{OV}(h) = \Psi\left(\mathcal{A}(h), w_v^h\right)$. Note that Ψ depends on the coalition structure $\mathcal{A}(h)$ but not on the network h directly. Consider the special case where the network $h \in \mathcal{H}$ actually forms a coalition structure: that is, assume $\bigcup_{c \in M} h(c) = N$ and $h(c) \cap h(c') \neq \varnothing$ only if $c = c'$ for all $c, c' \in M$. Phrased differently, each player is a member of exactly one connection. Here, the only difference between iteratively applying the Myerson Value IM and Owen's Value OV is that the latter one focuses less on single positions. More specifically, it is possible to show that OV satisfies the null player property only in a player-based way but not connection-wise. However, if the total productivity is determined by means of a component-additive value function, each player's contribution to the whole network is exactly what she contributes to her connection and therefore both axioms are equivalent.

Proposition 2.13. Assume $h \in \mathcal{H}$ forms a coalition structure and $v \in \mathcal{CA}$ is a component-additive value function. Then $v_i^{IM}(h) = v_i^{OV}(h)$ for all $i \in N$.[9]

The main idea of the proof is to translate the axiomatization given in Owen (1977) to the setting considered here and to show that both allocation rules have the same properties. The previous result immediately implies that the Position Value and Owen's Value differ by just one axiom under certain circumstances (i.e., if $h \in \mathcal{H}$ forms a coalition structure and $v \in \mathcal{V}$ is component-additive). Moreover, given these

[9]The author is grateful to André Casajus for pointing out a mistake in a previous version of this result.

requirements, it is straightforward to show that component-balancedness is equivalent to what Aumann and Drèze (1974) call "relative efficiency": The aggregated payoff of players within each connection equals exactly the connection's value. For this special case it is also possible to extend the Aumann-Drèze Value and it coincides with iteratively applying the Myerson Value and with Owen's Value.

2.5 Conclusion

The structures of many economical and social organizations are too complex to analyze them by means of Cooperative Game Theory or Bilateral Network Theory. The model introduced here uses mathematical hypergraphs (which are called coalitional networks in this study) to extend both approaches to a more general and richer framework. This enlarges not only the field of possible applications but it also allows to model them more realistically. In particular, the extended setting allows to formalize and analyze two-stage allocation procedures in a convenient way. Although these procedures already have been attempted in literature by some studies (e.g., Albizuri *et al.*, 2005; Aumann and Drèze, 1974; Borm *et al.*, 1992; Hart and Kurz, 1983; Owen, 1977) there has been no general framework for analyzing them in a structured way. This study not only provides such a framework but it also extends some particular examples of two-stage allocation rules well-known from Cooperative Game Theory and Network Theory to coalitional networks. Moreover, characterizations of these allocation rules are discussed. Interestingly, under certain requirements the Position Value and Owen's Value satisfy almost the same axioms.
The analysis of coalitional networks could be extended in several ways. One of the most interesting ones would be to investigate dynamic models of network formation. In the study presented here, the coalitional network was always given exogenously. Therefore, it seems to be crucial to know under which circumstances which structures will arise. To this end, suitable stability concepts and deviation rules would need to be introduced. It would furthermore be reasonable to allow N and M to be infinite and to work with finite carriers in order to have no exogenously given limitation in the number of networks. Another natural extension is to apply the extended model to applications not covered so far. Coalitional networks are richer than coalition structures or bilateral networks, and they allow to study settings where the players cooperate in quite complex structures. Examples of this are large organizations or social clubs, to name but a few.

Acknowledgements

The author is indebted to Sonja Brangewitz, André Casajus, and Claus-Jochen Haake for helpful comments and suggestions. He also thanks session participants at the workshop "Markets, Games & Design 2012 (MGD12)" in Bielefeld and Paderborn, at the conference "Public Economic Theory 2013 (PET13)" in Lisbon, and at the "9th Spain-Italy-Netherlands Meeting on Game Theory (SING9)" in Vigo. The author gratefully acknowledges financial support by the German Research Foundation (DFG) within the Collaborative Research Center "On-The-Fly Computing" (SFB 901).

2.6 Appendix: Proofs

Some of the proofs are based on only slightly varied but similar technical arguments. For example, the characterizations of the allocation rules discussed in this paper are all based on basic value functions and they all proceed in a similar way. In order to not repeat the same arguments again and again, the following two lemmas will be helpful.

Lemma 2.6. Let an allocation problem (N, M) be given. Assume the allocation rule $X \in \mathcal{X}$ satisfies (ADD), (IPLY), and (SYM). Furthermore, let $\alpha \in \mathbb{R}$ be a constant and $1^h \in \mathcal{V}$ a basic value function with $h \in \mathcal{H} \setminus \{h^{\varnothing}\}$. Then:

$$\left(X_i\left(\alpha 1^h\right)\right)(h') = \begin{cases} 0, & \text{if } i \notin N(h') \\ 0, & \text{if } i \in N(h') \text{ and } h \nsubseteq h' \\ \frac{\alpha}{|N(h')|}, & \text{if } i \in N(h') \text{ and } h = h' \end{cases}$$

Proof. Let $X \in \mathcal{X}$ satisfy (ADD), (IPLY), and (SYM). Furthermore, let $\alpha \in \mathbb{R}$, $1^h \in \mathcal{V}$ for $h \in \mathcal{H} \setminus \{h^{\varnothing}\}$, and $h' \in \mathcal{H}$. First note that the inactive players axiom implies $(X_i(\alpha 1^h))(h') = 0$ for all $i \notin N(h')$.

Case 1: $h \nsubseteq h'$.

This implies $1^h(h') = 0$. Furthermore, $\alpha 1^h(h'|_{S+i}) = \alpha 1^h(h'|_{S+j}) = 0$ for all $i, j \in N(h')$ and $S \subseteq N(h') \setminus \{i, j\}$. Thus, by symmetry $(X_i(\alpha 1^h))(h') = (X_j(\alpha 1^H))(h')$ for all $i, j \in N(h')$ and, because X is balanced by definition, this implies that all players receive a payoff of 0.

Case 2: $h = h'$.

This case proceeds analogously to the previous one. Again $\alpha 1^h(h'|_{S+i}) = \alpha 1^h(h'|_{S+j}) = 0$ for all $i,j \in N(h')$ and $S \subseteq N(h') \setminus \{i,j\}$. From this follows $(X_i(\alpha 1^h))(h') = (X_j(\alpha 1^h))(h')$ for all $i,j \in N(h')$ by symmetry and, thus, $\big(X_i\big(\alpha 1^h\big)\big)(h') = \frac{\alpha}{|N(h')|}$ by balancedness. □

Note that Lemma 2.6 makes no statement about $h \subsetneq h'$. For this case, further requirements are needed.

Lemma 2.7. Assume that the position-based allocation rule $Y \in \mathcal{Y}$ is additive. If it additionally satisfies (cwNULI) or (cwNULL), then $Y \in \mathcal{Y}^{2stg}$.

Proof. Let (N, M) be an arbitrary allocation problem. If $Y \in \mathcal{Y}$ is additive, for applying Proposition 2.3 it is sufficient to show that $v_c^Y = 0$ for $c \in M$ and $v \in \mathcal{V}$ always implies $v_{ic}^Y = 0$ for all $i \in N$. Phrased differently, if the connection's worth is zero in all networks the corresponding positions receive this worth as well. Therefore, suppose there is given a value function $v \in \mathcal{V}$ with $v_c^Y(h^*) = 0$ for all $h^* \in \mathcal{H}^*$. By applying either (cwNULI) or (cwNULL) this immediately implies $v_{ic}^Y(h) = 0$ for all $i \in N$. Thus, $Y \in \mathcal{Y}^{2stg}$. □

Proof of Proposition 2.2

This proposition is a direct implication of balancedness. Given an allocation problem (N, M) and a player-based allocation rule $Z \in \mathcal{X}$, each value function $v \in \mathcal{V}$ can be decomposed via $v = \sum_{c \in M} v_c^Z$. Thus, an equivalent formulation of invariance under redistribution for $X \in \mathcal{X}$ is

$$X_i\left(\sum_{c \in M} v_c^Z\right) = \sum_{c \in M} X_i\left(v_c^Z\right)$$

for all $i \in N$, $h \in \mathcal{H}_{M,2^M}$ and each sequence of payoff functions $\left\{v_c^Z\right\}_{c \in M}$, and this is obviously equivalent to additivity. □

Proof of Proposition 2.3

Let an allocation problem (N, M) be given. First assume $Y_{ic} = X_i \circ Z_c$ for two player-based allocation rules $X \in \mathcal{X}$ and $Z \in \mathcal{X}$. If $v_c^Y = \bar{v}_{\bar{c}}^Y$ for $c, \bar{c} \in M$ and $v, \bar{v} \in \mathcal{V}$, Lemma 2.1 implies

$$v_{ic}^Y = X_i\left(v_c^Z\right) = X_i\left(v_c^Y\right) = X_i\left(\bar{v}_{\bar{c}}^Y\right) = X_i\left(\bar{v}_{\bar{c}}^Z\right) = \bar{v}_{i\bar{c}}^Y \quad \text{for all } i \in N.$$

For the other direction let $Y \in \mathcal{Y}$ be given. According to Lemma 2.1 it is feasible to choose $Z_c := Z_c^Y$ for all $c \in M$. Exploiting the condition given in the proposition yields that

$$X_i(v') := \begin{cases} Y_{ic}(v), & \text{if there exist } c \in M \text{ and } v \in \mathcal{V} \text{ with } v' = Z_c(v) \\ X_i^Y(v'), & \text{otherwise} \end{cases}$$

is well-defined for all $v' \in \mathcal{V}$. Of course, it would also be possible to choose any other $X' \in \mathcal{X}$ in the second case. Then balancedness is indeed satisfied because

$$\begin{aligned} \sum_{i \in N} X_i(v') &= \begin{cases} \sum_{i \in N} Y_{ic}(v), & \text{if there exist } c \in M \text{ and } v \in \mathcal{V} \text{ with } v' = Z_c(v) \\ \sum_{i \in N} X_i^Y(v'), & \text{otherwise} \end{cases} \\ &= \begin{cases} Z_c^Y(v) = Z_c(v) = v', & \text{if there exist } c \in M \text{ and } v \in \mathcal{V} \text{ with } v' = Z_c(v) \\ \sum_{i \in N, c \in M} Y_{ic}(v') = v', & \text{otherwise.} \end{cases} \end{aligned}$$

Furthermore, by construction also $Y_{ic} = X_i \circ Z_c$ for all $i \in N$ and $c \in M$. □

Proof of Proposition 2.4

The proof proceeds analogously to the one of Proposition 2.3. □

Proof of Proposition 2.5

Let an allocation problem (N, M) be given. It is easy to verify that ED satisfies all the axioms. For the other direction let $\alpha \in \mathbb{R}$ and $h \in \mathcal{H}$. Furthermore, assume $X \in \mathcal{X}$ satisfies all the axioms. Let $h' \in \mathcal{H}$ be a further network. Applying Lemma 2.6 yields for $h \nsubseteq h'$ and $h = h'$ that $(X_i(\alpha 1^h))(h')$ is uniquely determined. Therefore, suppose for the rest of the proof $h \subsetneq h'$. Obviously, $\alpha 1^h = \alpha \sum_{\bar{h}: h \subseteq \bar{h}} 1_{\bar{h}}$, where $1_{\bar{h}}$ is called *Standard Value Function* (cf. van den Brink, 2007) and it is given by

$$1_{\bar{h}}(h') = \begin{cases} 1, & \text{if } \bar{h} = h' \\ 0, & \text{if } \bar{h} \neq h'. \end{cases}$$

If $\bar{h} \neq h'$, player $i \in N(h')$ is a nullifying player with respect to $\alpha 1_{\bar{h}}$ if and only if $i \notin N(\bar{h})$. This implies $(X_i(1_{\bar{h}}))(h') = 0$ for all $i \notin N(\bar{h})$. Furthermore, because of symmetry and balancedness the other players receive a payoff of zero, too. Thus, additivity implies

$$\left(X_i(\alpha 1^h)\right)(h') = \sum_{\bar{h}: h \subseteq \bar{h}} \underbrace{\left(X_i\left(\alpha 1_{\bar{h}}\right)\right)(h')}_{=0, \text{ if } \bar{h} \neq h'} = \left(X_i\left(\alpha 1_{h'}\right)\right)(h').$$

Analogously to case 2 in Lemma 2.6 it is possible to show that

$$(X_i(\alpha 1_{h'}))(h') = (X_i(\alpha 1^{h'}))(h')$$

and, therefore, the payoff is uniquely determined. □

Proof of Proposition 2.6

Let an allocation problem (N, M) be given. The proof proceeds in a similar way as the one of Proposition 2.5. First note that the Myerson Value satisfies the axioms. The easiest way for verifying this is by exploiting the properties of the Shapley Value Φ (see, e.g., Roth, 1988) because $v_i^{MV}(h) = \Phi_i(w_h^v)$, where $w_h^v(S) = v(h|_S)$ for all $S \subseteq N$. The rest of the proof proceeds in standard fashion. Because of additivity it is sufficient to concentrate on basic value functions for showing uniqueness. To this end let $\alpha \in \mathbb{R}$ and $h \in \mathcal{H} \setminus \{h^{\varnothing}\}$. Furthermore, assume $X \in \mathcal{X}$ satisfies the axioms given in the proposition. Because the null player property implies the inactive players axiom, Lemma 2.6 can be applied. That is, for $h \nsubseteq \bar{h}$ and $h = \bar{h}$ the payoff is uniquely determined. For $h \subsetneq \bar{h}$ each player $i \notin N(h)$ is a null player and therefore receives a payoff of 0. Because all players in $N(h)$ are symmetric in $\bar{h}$ and Y is balanced, $(X_i(\alpha 1^h))(\bar{h}) = \frac{\alpha}{|N(h)|}$ for all $i \in N(h)$. Thus, in this case the payoff is uniquely determined as well. □

Proof of Lemma 2.3

Let an allocation problem (N, M) be given and suppose $X \in \mathcal{X}$ satisfies additivity and the null player property. Furthermore, let $h \in \mathcal{H}$ be a network and $A_i \in \mathcal{A}(h)$ the association containing player $i \in N$. For each component-additive value function $v \in \mathcal{CA}$ define the auxiliary function $v_{A_i} : \mathcal{H} \longrightarrow \mathbb{R}$ by $v_{A_i}(\bar{h}) = v(\bar{h}|_{A_i})$ for all $\bar{h} \in \mathcal{H}$. Then:

$$\begin{aligned} v(h|_{S+i}) - v_{A_i}(h|_{S+i}) &= \sum_{A \in \mathcal{A}(h)} v\left(h|_{((S+i) \cap A)}\right) - v\left(h|_{((S+i) \cap A_i)}\right) \\ &= \sum_{A \in \mathcal{A}(h)} v\left(h|_{((S) \cap A)}\right) - v\left(h|_{((S) \cap A_i)}\right) = v(h|_S) - v_{A_i}(h|_S) \end{aligned}$$

Therefore, player i is a null player with respect to $v - v_{A_i}$ in the network h. The null player property implies $0 = (X_i(v_{A_i} - v))(h)$ and, thus, by additivity $(X_i(v))(h) = (X_i(v_{A_i}))(h)$ for all $v \in \mathcal{CA}$. Obviously, each player $j \notin A_i$ is a null player with respect to v_{A_i}. Hence,

$$\sum_{j\in A_i}(X_j(v))(h)=\sum_{j\in A_i}(X_j(v_{A_i}))(h)$$
$$=\sum_{j\in N}(X_j(v_{A_i}))(h)=v_{A_i}(h)=v(h|_{A_i}).$$

□

Proof of Proposition 2.8

Let an arbitrary allocation problem (N,M) be given and assume $Y\in\mathcal{Y}$ satisfies all of the axioms given in the proposition. In particular, Lemma 2.7 yields $Y\in\mathcal{Y}^{2stg}$. In the following let $X^c\in\mathcal{X}$ for $c\in M$ and $Z\in\mathcal{X}$ be the player-based allocation rules with $Y_{ic}=X_i^c\circ X_c$ for all $i\in N$ and $c\in M$. Proposition 2.2 implies $Z_c=Z_c^Y$ for all $c\in M$. Moreover, since Z^Y satisfies additivity, symmetry of active players, and the null player property by assumption, Proposition 2.6 implies that Z has to be the Myerson Value.
The remainder of the proof proceeds in a similar way as the characterization of the Equal Division Solution. In fact, as the axioms already indicate, it will be shown that within each connection the players distribute their value equally. Because of additivity it is sufficient to concentrate on basic value functions. To this end let $\alpha\in\mathbb{R}$ and $h\in\mathcal{H}$. Note that the inactive positions axiom implies $(Y_{ic}(\alpha 1^h))(h')=0$ for all $h'\in\mathcal{H}$ with $i\notin h'(c)$.

Case 1: $h\nsubseteq h'$.

Lemma 2.6 yields

$$\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right)=\left(MV_c\left(\alpha 1^h\right)\right)\left(h'^*\right)=0$$

for all $c\in M$. As all players $i,j\in h'(c)$ are symmetric within c, all positions in this connection receive the same payoff, i.e., $(Y_{ic}(\alpha 1^h))(h')=(Y_{jc}(\alpha 1^H))(h')$ for all $i,j\in h'(c)$. Consequently, $\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right)=0$ implies that this payoff has to be zero for all players.

Case 2: $h=h'$.

Here, Lemma 2.6 can be applied again:

$$\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right)=\left(MV_c\left(\alpha 1^h\right)\right)\left(h'^*\right)=\begin{cases}\frac{\alpha}{|M(h^*)|}, & \text{if } c\in M\left(h^*\right)\\ 0, & \text{if } c\notin M\left(h^*\right)\end{cases}$$

Similar to case 1, for all $c \in M$ all players within the connection are symmetric and by (cwSYM) all positions in c get the same. Thus,

$$\left(Y_{ic}\left(\alpha 1^h\right)\right)(h') = \frac{\alpha}{|h(c)|\,|M(h^*)|} \quad \text{for all } i \in h(c).$$

As already mentioned before, if $i \notin h(c)$ the corresponding position receives a payoff of zero.

Case 3: $h \subsetneq h'$.

Again, the connections' values are given by

$$\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right) = \left(MV_c\left(\alpha 1^h\right)\right)\left(h'^*\right) = \begin{cases} \frac{\alpha}{|M(h^*)|}, & \text{if } c \in M(h^*) \\ 0, & \text{if } c \notin M(h^*). \end{cases}$$

For $c \in M\left(h'^*\right) \setminus M(h^*)$ all players $i, j \in h'(c)$ are symmetric within c and, similar to case 1, applying (cwSYM) yields

$$\left(Y_{ic}(1^h)\right)(h') = \left(Y_{jc}(1^h)\right)(h') = 0.$$

Therefore, suppose $c \in M(h^*)$. Furthermore, let again $\alpha 1^h = \alpha \sum_{\bar{h}: h \subseteq h'} 1_{\bar{h}}$ where the $1_{h'}$ are standard value functions (cf. the proof of Proposition 2.5). If there exists no $T \subseteq M$ with $\bar{h} = h'|_T$, then player $i \in h'(c)$ is a nullifying player in c with respect to $\alpha 1_{\bar{h}}$ if and only if $i \notin \bar{h}(c)$. Thus, $\left(Y_{ic}(\alpha 1_{\bar{h}})\right)(h') = 0$ for all $i \notin h'(c)$. Because each connection $c' \in M$ is a null player in h' with respect to $1_{\bar{h}}$, the aggregated value $\left(Z_c^Y(\alpha 1_{\bar{h}})\right)(h') = \sum_{i \in N}\left(Y_{ic}(\alpha 1_{\bar{h}})\right)(h')$ of c also has to be equal to zero and by (cwSYM) this implies $\left(Y_{ic}(\alpha 1_{\bar{h}})\right)(h') = 0$ for all $i \in N$. Thus, $\left(Y_{ic}(\alpha 1_{\bar{h}})\right)(h') \neq 0$ only if there exists $T \subseteq M$ with $\bar{h} = h'|_T$. Now the value of position $i \in h'(c)$ can be decomposed in the following way:

$$\begin{aligned} \left(Y_{ic}(\alpha 1^h)\right)(h') &= \sum_{\bar{h}: h \subseteq \bar{h}} \underbrace{\left(Y_{ic}\left(\alpha 1_{\bar{h}}\right)\right)(h')}_{=0, \text{ if } \nexists\, T \subseteq M:\, \bar{h} = h'|_T} \\ &= \sum_{T: M(h^*) \subseteq T \subseteq M(h'^*)} \left(Y_{ic}\left(\alpha 1_{h'|_T}\right)\right)(h') \qquad (2.2) \end{aligned}$$

Note that $c \in M(h^*)$ implies $c \in T$ for all $M(h^*) \subseteq T \subseteq M\left(h'^*\right)$ and, thus, $h'|_T(c) = h'(c)$. Therefore, all players in $h'(c)$ are symmetric in c with respect to $1_{h'|_T}$ and (cwSYM) yields

$$\left(Y_{ic}\left(1_{h'|_T}\right)\right)(h') = \frac{\left(Z_c^Y\left(\alpha 1_{h'|_T}\right)\right)\left(h'^*\right)}{|h'(c)|}$$

for all $i \in h'(c)$. Now Equation (2.2) can be exploited again:

$$\begin{aligned}\left(Y_{ic}(\alpha 1^h)\right)(h') &= \sum_{T:M(h^*)\subseteq T\subseteq M(h'^*)} \left(Y_{ic}\left(\alpha 1_{h'|_T}\right)\right)(h') \\ &= \frac{1}{|h'(c)|} \sum_{T:M(h^*)\subseteq T\subseteq M\left(h'^*\right)} \left(Z_c^Y\left(\alpha 1_{h'|_T}\right)\right)\left(h'^*\right) \\ &= \frac{1}{|h'(c)|} \sum_{\bar{h}:h\subseteq\bar{h}} \left(Z_c^Y\left(\alpha 1_{h'|_T}\right)\right)\left(h'^*\right) \\ &= \frac{1}{|h'(c)|}\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right) = \frac{\alpha}{|h'(c)|\left|M\left(h^*\right)\right|}\end{aligned}$$

From this it follows that within each connection the value is allocated equally among all corresponding members. □

Proof of Proposition 2.11

Let an arbitrary allocation problem (N, M) be given. Proposition 2.6 implies that iteratively applying the Myerson Value satisfies the axioms. For the other direction let $Y \in \mathcal{Y}$. Moroever, assume Y satisfies all the axioms. By applying Lemma 2.7 this implies $Y \in \mathcal{Y}^{2stg}$. First note that because of additivity it is again sufficient to concentrate on $\alpha 1^h$, where $\alpha \in \mathbb{R}$ and $h \in \mathcal{H}$.[10] Furthermore, let $h' \in \mathcal{H}$.

Case 1: $h \nsubseteq h'$.

Here, every player is a null player in each connection. Thus, $(Y_{ic}(\alpha 1^h))(h') = 0$ for all $i \in N$ and $c \in M$.

Case 2: $h \subseteq h'$.

In this case, ic is a null position if and only if $i \notin h(c)$. From this follows

$$\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right) = \sum_{i\in N}\left(Y_{ic}(\alpha 1^h)\right)(h') = 0 \quad \text{for all } c \in M \text{ with } |h'(c)| = 0$$

and, thus, symmetry of Z^Y implies $\left(Z_c^Y\left(\alpha 1^h\right)\right)\left(h'^*\right) = \frac{\alpha}{|M(h')|}$ for all $c \in M(h')$. Finally, exploiting (cwSYM) yields $(Y_{ic}(\alpha 1^h))(h') = \frac{\alpha}{|h(c)||M(h^*)|}$ for all $i \in N$ and $c \in M$ with $i \in h(c)$.

[10] Although Owen (1977) requires the null player property, the remainder of the proof in principle proceeds completely analogously to the characterization of his value. Of course, (NULL) has to be replaced by (cwNULL) but the line of argumentation is the same.

Therefore, Y equals iteratively applying the Myerson Value. In particular, according to Proposition 2.6 this implies that the first-stage allocation rule satisfies (NULL). □

Proof of Proposition 2.13

The main idea of the proof is to show that if a network forms a coalition structure and moreover a component-additive value function is given, IM and OV satisfy the same axioms. To this end, it is necessary to introduce some preliminaries from Cooperative Game Theory. As already mentioned in Remark 2.1, a TU game is a pair (N, γ) with $\gamma : 2^N \longrightarrow \mathbb{R}$ and $\gamma(\varnothing) = 0$. If the set of players is decomposed into a set of coalitions $B = \{B_1, \ldots, B_k\}$, i.e., $\bigcup_{l=1}^{k} B_l = N$ and $B_l \neq \varnothing$ for all $1 \leq l \leq k$, this set B is called *partition*. Moreover, a (single-valued) *solution* Γ assigns a payoff vector $\Gamma(\gamma, B) \in \mathbb{R}^n$ to each pair consisting of a TU game (N, γ) and a partition B. Analogously to allocation rules it is possible to characterize solutions by means of several axioms: Let two arbitrary TU games (N, γ) and (N, γ') together with a partition B be given. A solution Γ...

Efficiency ...is *efficient* if $\sum_{i \in N} \Gamma_i(\gamma, B) = \gamma(N)$.

Additivity ...is *additive* if $\Gamma(\gamma + \gamma', B) = \Gamma(\gamma, B) + \Gamma(\gamma', B)$.

Null player property ...satisfies the *null player property* if every null player $i \in N$ (i.e., $\gamma(S \cup \{i\}) - \gamma(S) = 0$ for all $S \subseteq N \smallsetminus \{i\}$) always receives a payoff of zero: $\Gamma_i(v, B) = 0$.

Symmetry within coalitions ...satisfies *symmetry within coalitions* if two symmetric players $i, j \in N$ (i.e., $\gamma(S \cup \{i\}) = \gamma(S \cup \{j\})$ for all $S \subseteq N \smallsetminus \{i, j\}$) always receive the same payoff: $\Gamma_i(v, B) = \Gamma_j(v, B)$.

Symmetry across coalitions ...satisfies *symmetry across coalitions* if the aggregated payoff of two symmetric coalitions $B_l, B_{l'} \in B$ (i.e., $v(B' \cup B_l) = v(B' \cup B_{l'})$ for all $B' \subseteq B \smallsetminus \{B_l, B_{l'}\}$) is always the same: $\sum_{i \in B_l} \Gamma_i(v, B) = \sum_{i \in B_l} \Gamma_i(v, B)$.

Owen (1977) has shown the following result:[11]

Theorem. Owen's value Ψ is the unique solution satisfying efficiency, additivity, the null player property, symmetry within coalitions, and symmetry across coalitions.

[11] Actually, Owen (1977) used slightly different but equivalent axioms in his work. The more common variation of his axiomatization stated here can be found, for example, in Winter (2002).

Now, let an allocation problem (N, M) and a component-additive allocation rule $v \in \mathcal{CA}$ be given. Moreover, suppose the network $h \in \mathcal{H}$ forms a coalition structure. Then, each player-based allocation rule $X \in \mathcal{X}$ can be interpreted as a position-based allocation rule via

$$v_{ic}^X(h) := \begin{cases} v_i^X(h), & \text{for } i \in h(c) \\ 0, & \text{for } i \notin h(c). \end{cases}$$

According to Corollary 2.4 from Hart and Kurz (1983), $(Z_c^{OV})_{c \in M}$ equals the Myerson Value:

$$v_c^{OV}(h^*) = \sum_{i \in N} v_{ic}^{OV}(h) = \sum_{i \in h(c)} v_i^{OV}(h) = \sum_{i \in h(c)} \Psi_i\left(\mathcal{A}(h), w_v^h\right) \overset{\text{HK}}{=} \Phi_c\left(w_v^{h^*}\right) = v_c^{MV}(h^*),$$

where, again, Φ denotes the Shapley Value and $w_v^{h^*}(T) := v(h|_T^*)$ for all $T \subseteq M$ (cf. Footnote 6). In particular, this implies $(Z_c^{OV})_{c \in M}$ satisfies symmetry. By applying the characterization of Ψ it is straightforward to show that OV also satisfies additivity and the null player property (with respect to coalitional networks). Because h is a coalition structure and v is component-additive, player i's contribution to the whole network is exactly what she contributes to her connection $c \in M$:

$$\begin{aligned} v(h|_{S \cup \{i\}}) - v(h|_S) &= v(h|_{h(c) \cap (S \cup \{i\})}) - v(h|_{h(c) \cap S}) \\ &= v_c^{MV}\left(h|_{S \cup \{i\}}^*\right) - v_c^{MV}\left(h|_S^*\right) \\ &= v_c^{MV}\left(h|_{c, S \cup \{i\}}^*\right) - v_c^{MV}\left(h|_{c,S}^*\right) \end{aligned}$$

for all $S \subseteq N$. Thus, OV satisfies the null player property not only in a player-based way but also connection-wise. Nevertheless, it should be mentioned that this is true only under the requirements given here. By further exploiting the construction of the Myerson Value and component additivity, it is possible to show that symmetry within coalitions corresponds to connection-wise symmetry (note that from Lemma 2.7 it follows that $OV \in \mathcal{Y}^{2stg}$): Let $i, j \in N$ such that there exists $c \in M$ with $i, j \in h(c)$. Then:

$$\begin{aligned} & & v(h|_{S \cup \{i\}}) &= v(h|_{S \cup \{j\}}) & & \text{for all } S \subseteq N \setminus \{i, j\} \\ &\Leftrightarrow & \sum_{c' \in M} v(h|_{h(c') \cap (S \cup \{i\})}) &= \sum_{c' \in M} v(h|_{h(c') \cap (S \cup \{j\})}) & & \text{for all } S \subseteq N \setminus \{i, j\} \\ &\Leftrightarrow & v(h|_{c, S' \cup \{i\}}) &= v(h|_{c, S' \cup \{j\}}) & & \text{for all } S' \subseteq h(c) \setminus \{i, j\} \\ &\Leftrightarrow & v(h|_{c, S' \cup \{i\}}^*) &= v(h|_{c, S' \cup \{j\}}^*) & & \text{for all } S' \subseteq h(c) \setminus \{i, j\} \\ &\Leftrightarrow & v_c^{MV}(h|_{c, S' \cup \{i\}}^*) &= v_c^{MV}(h|_{c, S' \cup \{j\}}^*) & & \text{for all } S' \subseteq h(c) \setminus \{i, j\} \end{aligned}$$

Therefore, OV and IM satisfy the same axioms. □

References

Albizuri, M. J., Aurrecoechea, J. and Zarzuelo, J. (2005). Configuration values: Extensions of the coalitional Owen value. *Games and Economic Behavior*, **57**, 1–17.

Arenas, A., Cabrales, A., Danon, L., Diaz-Guilera, A., Guimerà, R. and Vega-Redondo, F. (2010). Optimal information transmission in organizations: search and congestion. *Review of Economic Design*, **14**, 75–93.

Aumann, R. J. and Drèze, J. H. (1974). Cooperative games with coalition structures. *International Journal of Game Theory*, **3** (4), 217–237.

Berge, C. (1989). *Hypergraphs: Combinatorics of Finite Sets*. North Holland.

Bolton, P. and Dewatripont, M. (1994). The firm as a communication network. *The Quarterly Journal of Economics*, **109** (4), 809–839.

Borm, P., Owen, G. and Tijs, S. (1992). On the position value for communication situations. *SIAM Journal on Discrete Mathematics*, **5** (3), 305–320.

Caulier, J., Mauleon, A. and Vannetelbosch, V. (2012). Contractually stable networks. *International Journal of Game Theory*, **Forthcoming**.

Chalkiadakis, G., Elkind, E., Markakis, E., Polukarov, M. and Jennings, N. (2010). Cooperative games with overlapping coalitions. *Journal of Artificial Intelligence Research*, **39**, 179–216.

Deegan, J. and Packel, E. W. (1978). A new index of power for simple n-person games. *lnternational Journal of Game Theory*, **7** (2), 113–123.

Fershtman, C. and Persitz, D. (2012). Clubs and social networks. *Mimeo*.

Harsanyi, J. C. (1959). A bargaining model for the cooperative n-person game. In A. Tucker (ed.), *Contributions to the Theory of Games IV*, Princeton: Princeton University Press, pp. 325–355.

Hart, S. and Kurz, M. (1983). Endogenous formation of coalitions. *Econometrica*, **51** (4), 1047–1064.

Jackson, M. O. (2005). Allocation rules for network games. *Games and Economic Behavior*, **51**, 128–154.

— (2008). *Social and Economic Networks*. Princeton University Press.

— and Wolinsky, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, **71** (0108), 44–74.

Jorzik, N. (2012). Allocation rules for hypergraph games. *Mimeo*.

Myerson, R. B. (1977). Maximin facility location. *Journal of Research*, **79** (B), 107–115.

— (1980). Conference structures and fair allocation rules. *International Journal of Game Theory*, **9** (3), 169–182.

Owen, G. (1977). Values of games with a priori unions. In R. Hein and O. Moeschlin (eds.), *Essays in Mathematical Economics and Game Theory*, New York: Springer Verlag, pp. 76–88.

Radner, R. (1993). The organization of decentralized information processing. *Econometrica*, **61** (5), 1109–1146.

Roth, A. (ed.) (1988). *The Shapley Value*, Cambridge: Cambridge University Press.

Shapley, L. S. (1953). A value for n-person games. In H. W. Khun and A. W. Tucker (eds.), *Contributions to the Theory of Games*, Princeton: Princeton University Press, pp. 307–314.

Sotomayor, M. (2004). Implementation in the many-to-many matching market. *Games and Economic Behavior*, **46**, 199–212.

van den Brink, R. (2007). Null or nullifying players: The difference between the shapley value and equal division solutions. *Journal of Economic Theory*, **136**, 767–775.

van den Nouweland, A., Borm, P. and Tijs, S. (1992). Allocation rules for hypergraph communication situations. *International Journal of Game Theory*, **20** (3), 255–268.

— and Slikker, M. (2012). An axiomatic characterization of the position value for network situations. *Mathematical Social Sciences*, **Forthcoming**.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis*. Cambridge University Press.

Winter, E. (1989). A value for cooperative games with levels structure of cooperation. *International Journal of Game Theory*, **18**, 227–240.

— (2002). The shapley value. In R. J. Aumann and S. Hart (eds.), *The Handbook of Game Theory with Economic Applications*, Amsterdam: North-Holland, vol. 3, pp. 2025–2054.

Chapter 3

Constitutions and Social Networks

3.1 Introduction

There are various situations in economic or daily life where individuals organize themselves in groups, whether for cooperation, coordination, or otherwise. The goal of this paper is to formalize and examine environments where individuals are allowed to engage in several groups at the same time. These group structures are interpreted as social networks in this study. Depending on the context, formation of these networks occurs for manifold reasons and considering all of them seems to be a virtually impossible venture. In order to be as general as possible, we abstract from activities carried out within each group. That is, we suppose the individuals' preferences to depend on the structure of the network directly. Given these preferences, there might be incentives for joining or leaving certain groups. The salient point is, however, that individuals are generically not necessarily free to deviate. Some members of a group might have certain property rights which allow them to block new members or even give them the power to force existing members to stay. We capture this aspect by means of introducing the notion of constitutions. More precisely, each group is supposed to have specific rules governing, on the one hand, which deviations are feasible and, on the other, who may decide about the deviations. Therefore, the formation of social networks not only depends on the preferences of the individuals but also on the property rights granted by the constitutions.

The framework outlined above captivates with a wide spectrum of possible applications. A particular one that we are going to discuss in detail is job markets with labor unions. But one could also mention research collaborations, immigration, or social clubs, for instance. These examples already indicate that the rules or constitutions governing which members may join or leave a group may vary greatly. For instance, in some groups it might be possible to dismiss members but in others there might be a protection against this. Or, in some groups entry might be free but in others it might require the consent of other members. Therefore, the constitutional design may have a significant impact on the formation of social networks. Consequently, two questions which we are going to address are: (i) what changes if more blocking power is given to the individuals and, more general, (ii) under which circumstances is it possible to find constitutions which guarantee a certain degree of stability?

Since the formation of social groups is of fundamental interest, it has been examined from numerous angles before. For instance, Ellickson *et al.* (1999, 2001) as well as Allouch and Wooders (2008) analyze this issue in the framework of general equilibrium theory, Acemoglu *et al.* (2012) provide a dynamic model for studying the stability of societies, and Page and Wooders (2010) formalize club formation as a non-cooperative game, to name but a few. In fact, providing a complete overview over all publications dealing with group formation in a broader sense would exceed the scope of nearly every paper due to the great complexity and diversification of the field. Therefore, the following survey restricts on most closely related branches and outlines which publications particularly influenced our work.

Analyzing group formation but abstracting from activities carried out within each group obviously relates to hedonic coalition formation (e.g., Banerjee *et al.*, 2001; Bogomolnaia and Jackson, 2002). Moreover, studies dealing with economic networks (e.g., Jackson, 2008) or matching markets (e.g., Roth and Sotomayor, 1990) can also be embedded into our setting. Thus, we contribute indirectly to a stream of literature where the authors combine coalition formation and matching problems (e.g., Cesco, 2012; Pycia, 2012). However, the way we model social networks and preferences is closer to models from matching theory where the individuals are not only concerned about which groups they belong to but also about who the other members of the groups are (e.g., Dutta and Massó, 1997; Eichenique and Yenmez, 2007; Kominers, 2010).

One of the main contributions of this paper is formalizing constitutional rules within a hedonic setting. This approach is in spirit with some other publications from literature, like Bala and Goyal (2000), Page and Wooders (2009), or Jehiel and Scotchmer (2001), for example. These papers analyze which networks or coalition structures might be expected to emerge under several specific rules governing network or coalition formation, respectively. However, the aforementioned works differ from ours in at least one important aspect: For analyzing which structures are likely to occur we focus on constitutionally stable networks, where a social network is considered to be constitutionally stable if none of the groups is modified any more. The salient point is that, in our framework, the stability of a network depends on explicitly modeled constitutions. In the above-mentioned papers, on the contrary, the constitutional rules are varied only implicitly by discussing different stability concepts. For this reason, our approach not only achieves greater generality but it also allows separating more clearly which influence constitutional rules have on group formation.

The analysis conducted in this paper is two-fold: We not only focus on the question whether constitutionally stable networks actually exist but we also discuss whether they might be reached given that the players apply a "trial-and-error strategy". To this end, we follow Roth and Vande Vate (1990). In the context of marriage problems (or two-sided one-to-one matchings, respectively), the authors introduced a Markov process which always results in a stable matching with probability one, even if the individuals act myopically. Later, this work has been extended and varied in several ways (e.g., Chung, 2000; Diamantoudi *et al.*, 2004; Klaus *et al.*, 2010; Kojima and Ünver, 2008). In our study, we use basically the same approach but we adopt the terminology of Jackson and Watts (2001, 2002) who examined a similar random process but focused on stochastic stability of economic networks. By analyzing "improving paths" we formulate requirements on constitutions and preferences guaranteeing that for every social network there always exists an improving path leading to a stable network. It turns out, in fact, that this is equivalent to requiring the existence of a specific version of a common ranking (cf. Banerjee *et al.*, 2001; Farrell and Scotchmer, 1988). We also find that giving more blocking power to the individuals does not necessarily lead to more stability. Indeed, higher blocking power might destroy the existence of the common ranking.

Although the main purpose of this paper is to discuss formation of social networks in general, the last part is devoted to a particular application, namely to job markets with labor unions. Applying the general results obtained in the sections before allows to judge for different levels of the unions' strength, whether the job market is likely to become stable or not. In doing so, we also find a variation of Roth's "polarization of interests" (cf. Roth, 1984) between employers and employees.

The remainder of the paper proceeds as follows: The next section introduces the model. This also includes formal definitions of social networks and of constitutions. In Section 3.3, we discuss conditions for the existence of constitutionally stable networks and, in Section 3.4, we apply the corresponding results to our model of job markets. Finally, Section 3.5 contains the conclusions.

3.2 The Model

In the following, let $N = \{i_1, \ldots, i_n\}$ be a finite set of *players* and $M = \{c_1, \ldots, c_m\}$ be a finite set of *connections*.

Definition 3.1. A *social network* h is a mapping $h : M \longrightarrow 2^N$ assigning to each $c \in M$ a subset of players (as usual, 2^N is the power set of N).[1]

A social network h indicates which players are members of which connections. For each $i \in N$ let $M_h(i) = \{c \in M \mid i \in h(c)\}$ be the set of connections player i is contained in. The set of all social networks is denoted by $\mathcal{H}$ and the cardinality of $\mathcal{H}$ is $|\mathcal{H}| = 2^{mn}$. A particular special case is the *empty social network* $h^{\varnothing} \in \mathcal{H}$, with $h^{\varnothing}(c) = \varnothing$ for all $c \in M$. That is, no player is contained in any connection.

Example 3.1. *Suppose there are three players and four connections, i.e., we have $N = \{i_1, i_2, i_3\}$ and $M = \{c_1, \ldots, c_4\}$. Consider the case where all players are contained in c_1, the players i_2 and i_3 are moreover contained in c_2 and c_3, while c_4 only contains i_1. This is described formally by the following social network h:*

$$h(c) = \begin{cases} \{i_1, i_2, i_3\}, & \text{if } c = c_1 \\ \{i_2, i_3\}, & \text{if } c \in \{c_2, c_3\} \\ \{i_1\}, & \text{if } c = c_4. \end{cases}$$

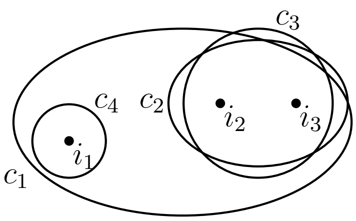


Figure 3.1: The social network h

3.2.1 Constitutions

Each player $i \in N$ is supposed to have rational preferences $\succeq^i$ over $\mathcal{H}$ and the tuple $\succeq = (\succeq^i)_{i \in N}$ is called a *preference profile*. Depending on the preferences, there might

[1] Note that the tuple (N, M, h) is simply a mathematical hypergraph. Therefore, from a technical point of view our definition of social networks also relates to the notions of conference structures (e.g., Myerson, 1980) and many-to-many matchings (e.g., Roth, 1984).

be incentives to alter some connection in a given network. For modeling this formally we use the symmetric difference $\pm$ defined by $D' \pm D = (D' \setminus D) \cup (D \setminus D')$ for all $D', D \subseteq N$. Correspondingly, given a connection $c \in M$ and a subset of players $D \subseteq N$, let $h \pm (c, D)$ be the social network that is obtained from $h \in \mathcal{H}$ if c is altered by the players in D. More specifically, players in $D \cap h(c)$ leave the connection and players in $D \setminus h(c)$ join it:[2]

$$(h \pm (c, D))(c') := \begin{cases} h(c) \pm D & \text{if } c = c' \\ h(c') & \text{if } c \neq c' \end{cases} \tag{3.1}$$

If $D \cap h(c) = \varnothing$, we just write $h + (c, D)$ instead of $h \pm (c, D)$ to stress the fact that no player leaves the connection. Conversely, if $D \subseteq h(c)$, we just write $h - (c, D)$ instead of $h \pm (c, D)$ to indicate that no player joins the connection.
The central assumption in our framework is that certain deviations might be precluded, even if all deviating players would benefit from altering the network. For capturing this facet, we introduce constitutions which govern the exit of already existing members and/or the arrival of new members. That is, the constitutions describe, on the one hand, which modifications of a connection are feasible and, on the other, the coalitions whose support is needed for the modifications to take place.

Definition 3.2. The *constitution* $\mathcal{C}^c = (\mathcal{C}^c_h)_{h \in \mathcal{H}}$ of connection $c \in M$ is a collection of pairs $\mathcal{C}^c_h = (\mathcal{D}^c_h, S^c_h)$ where (i) $\mathcal{D}^c_h \subseteq 2^N \setminus \{\varnothing\}$ describes the *feasible deviations* and (ii) for each $D \in \mathcal{D}^c_h$, $S^c_h(D) \subseteq 2^{h(c)}$ specifies a non-empty set of *supporting coalitions*.

For all $c \in M$ and $h \in \mathcal{H}$, $\mathcal{C}^c_h$ consists of two components. The first one, $\mathcal{D}^c_h$, specifies which changes (with respect to the deviations formalized in (3.1)) of c are possible. Of course, it might be the case that $\mathcal{D}^c_h = 2^N \setminus \{\varnothing\}$ and that there are no restrictions on feasible deviations. In many applications, however, certain modifications of a connection are not possible due to capacity constraints or legal requirements, for example, and this is captured by $\mathcal{D}^c_h$. Moreover, for deviating from $h(c)$, each deviating set of players $D \in \mathcal{D}^c_h$ needs the support of at least one *supporting coalition* $S \in S^c_h(D)$.[3] If there exists no such S, the modification by D is blocked. Note that

[2]We use $\pm$ instead of the usual symbol Δ for denoting the symmetric difference in order to emphasize that it might be possible that at the same time new members enter a connection while other members leave it.

[3]Our notions of feasible deviations and supporting coalitions relate in a way to "move arcs" and "preference arcs", respectively, from Page *et al.* (2005).

$\varnothing \in S_h^c(D)$ is allowed, too. In this case, the players in $D \in \mathcal{D}_h^c$ do not need the consent of any member of the connection for deviating. Moreover, if $S \in \mathcal{S}_h^c(D) \setminus \{\varnothing\}$, we assume $S' \in \mathcal{S}_h^c(D)$ for all $S' \supseteq S$. That is, if S is a supporting coalition for a certain deviation, all coalitions containing S also have the power to support this deviation. In the following, let $\mathcal{C} := (\mathcal{C}^c)_{c \in M}$. The tuple $(N, M, \succeq, \mathcal{C})$ is called a *society*.

Example 3.2. *Let $N = \{i_1, \ldots, i_n\}$ and $M = \{c_1, c_2, c_3\}$. As an example consider the following three specific constitutions:*

(i) If $\mathcal{D}_h^{c_1} = \{D \subseteq N \,|\, |h(c_1) \pm D| \leq 9,\, D \neq \varnothing\}$ and $\mathcal{S}_h^{c_1}(D) = \{S \subseteq h(c_1) \,|\, 2 \cdot |S| > |h(c_1)|\}$ for all $h \in \mathcal{H}$ and $D \in \mathcal{D}_h^{c_1}$, the players have to respect a quota of nine and decisions are taken by means of the majority rule.

(ii) Suppose $\mathcal{D}_h^{c_2} = \{D \subseteq N \,|\, l \geq 3 \;\forall\, i_l \in D,\, D \neq \varnothing\}$ and $\mathcal{S}_h^{c_2}(D) = \{S \subseteq h(c_2) \,|\, h(c_2) \cap D \subseteq S\}$ for all $h \in \mathcal{H}$ and $D \in \mathcal{D}_h^{c_1}$. This reflects the case where deviations require certain qualifications. In this specific example, players need an index of at least three. Moreover, none of the members has property rights for the connection. If a deviation is feasible, the corresponding players have the power to support themselves, i.e., they are free to enter or exit.

(iii) Let $\mathcal{D}_h^{c_3} = 2^N \setminus \{\varnothing\}$ and $\mathcal{S}_h^{c_3}(D) = \{S \subseteq h(c_3) \,|\, i_{\bar{l}} \in S,\ \text{where } \bar{l} \geq l \;\forall\, i_l \in h(c_3)\}$ for all $h \in \mathcal{H}$ and $D \in \mathcal{D}_h^{c_1}$. Here, all deviations are feasible and the player with the highest index acts as a kind of dictator and has perfect property rights. That is, she may decide about both, whether players may join the connection as well as whether they may leave it.

3.2.2 Stability

For analyzing which social networks might be expected to emerge we propose a notion of stability which requires that no single connection is altered any more.

Definition 3.3. Given the society $(N, M, \succeq, \mathcal{C})$, a social network h is *constitutionally stable* with respect to the constitutions $\mathcal{C}$ if for all $c \in M$ and $D \in \mathcal{D}_h^c$ we have that: (i) $h \succeq^i h \pm (c, D)$ for at least one $i \in D \setminus h(c)$ or (ii) in each supporting coalition $S \in \mathcal{S}_h^c(D)$ there is a player $j \in S$ with $h \succeq^j h \pm (c, D)$.

Expressed in words, a social network $h \in \mathcal{H}$ is constitutionally stable if and only if for any connection $c \in M$ and any feasible modification $D \in \mathcal{D}_h^c$, at least one of the players joining c does not strictly benefit from deviating or at least one of the members of every supporting coalition $S \in \mathcal{S}_h^c(D)$ is not strictly better off. Therefore, implicitly we assume that moving from $h \in \mathcal{H}$ to $h \pm (c, D)$ does not necessarily need the consent of players leaving c. The main idea is that some members of the connection might have the power to force other members to leave c even when the excluded players suffer from this deviation. On the other hand, a player who is not in $c \in M$ cannot be forced to join the connection. Only if she strictly benefits will she agree to becoming a member of it.

Remark 3.1. An alternative approach one might think about would be that a deviation already takes place if only one ot the supporting or deviating players is strictly better off and the others do not suffer from the deviation. This would be a direct extension of "pairwise stability" from Network Theory (see Jackson and Wolinsky, 1996). However, in our model this variation causes strange curiosities which are not plausible in real life. For example, consider the following situation: Suppose there are three players $N = \{i_1, i_2, i_3\}$ and only one connection $M = \{c\}$. Let $h(c) = \{i_1, i_2\}$, $h'(c) = \{i_1, i_3\}$, $\{i_2, i_3\} \in \mathcal{D}_h^c \cap \mathcal{D}_{h'}^c$, and $\{i_1\} \in \mathcal{S}_h^c(\{i_2, i_3\}) \cap \mathcal{S}_{h'}^c(\{i_2, i_3\})$. Moreover, suppose the preferences with respect to these two networks are as follows: $h \sim^{i_1} h'$, $h \succ^{i_2} h'$, and $h' \succ^{i_3} h$. Then, given the network h, if player i_2 is replaced by i_3, the last-mentioned is strictly better off while i_2 suffers from this deviation. Nevertheless, since i_1 is indifferent between h and h', she would support deviating from h to h' because she has no incentive for blocking this modification of the connection. However, in the next step, the same pattern recurs again. The only difference is that now i_2 and i_3 interchange roles. Thus, i_1 would grant i_2 access to the connection although she evicted this player in the step before. In particular, this skipping back and forth between the two networks the whole time implies an inconsistency in the behavior of i_1 which is quite counterintuitive. For excluding situations like these, we require that players only deviate or support a deviation if they are strictly better off. Note that this is in line with several other related stability concepts from literature, like the core stability of Bogomolnaia and Jackson (2002) and Banerjee *et al.* (2001), the pairwise stability of Sotomayor (1999), or the strong stability of Dutta and Mutuswami (1997), to name but a few.

In the following, let $\mathcal{ST}(\mathcal{C})$ denote the set of constitutionally stable networks with respect to the constitutions $\mathcal{C}$. Moreover, for each $h \in \mathcal{H}$, let

$$\mathcal{A}_h^c(\mathcal{C}) := \{D \in \mathcal{D}_h^c \mid \exists\, S \in \mathcal{S}_h^c(D) \text{ such that } h \pm (c, D) \succ^i h \;\forall\, i \in (D \setminus h(c)) \cup S\}$$

be the set of all feasible deviations causing instabilities in $c \in M$. Note that if $D \subseteq h(c)$, $D \in \mathcal{D}_h^c$, and $\varnothing \in \mathcal{S}_h^c(D)$, then D causes an instability by definition although it might be that nobody benefits from this deviation. Therefore, in order to exclude exogenous instabilities like these, we will assume $\varnothing \notin \mathcal{S}_h^c(D)$ if $D \subseteq h(c)$.

3.3 General Results

Generically constitutionally stable social networks might fail to exist and this leads to the question of how the design of constitutions affects the (non-)existence of stable structures. For approaching this issue let us start with a straightforward and plausible attempt: Suppose the constitutions grant the players a certain level of blocking power. That is, the members of each connection might have certain property rights allowing them to inhibit modifications of the connection which are not conform to their own ideas.

Remark 3.2. Let two societies $(N, M, \succeq, \mathcal{C})$ and $(N, M, \succeq, \bar{\mathcal{C}})$ be given and assume $\mathcal{C} \subseteq \bar{\mathcal{C}}$, i.e., $\mathcal{D}_h^c \subseteq \bar{\mathcal{D}}_h^c$ and $\mathcal{S}_h^c(D) \subseteq \bar{\mathcal{S}}_h^c(D)$ for all $h \in \mathcal{H}$, $c \in M$, and $D \in \mathcal{D}_h^c$. Then: $\mathcal{ST}(\bar{\mathcal{C}}) \subseteq \mathcal{ST}(\mathcal{C})$.

The remark follows directly from the definition of constitutional stability. If the sets of feasible deviations and supporting coalitions shrink, the blocking power of each individual player increases and the set of constitutionally stable networks might become larger. However, although the reasoning is very intuitive it might be misleading. In fact, whether more blocking power really implies more stability, strongly depends on the perspective of stability: On the one hand, there might be more stable networks but, on the other, reaching them might not be possible any more.
For formalizing these thoughts we follow Jackson and Watts (2001, 2002) and use the notion of improving paths: "An improving path is a sequence of networks that can emerge when individuals [join or leave a connection] based on the improvement the resulting network offers relative to the current network" (Jackson and Watts, 2002, p. 51). That is, each of the networks differs from its predecessor only in that exactly one connection is modified by a deviating coalition. This requires, of course, that every player joining the connection must strictly prefer the resulting network to the current one. Moreover, the deviation should not be blocked and, hence, there should be a supporting coalition where every member strictly benefits from the modification.[4] More formally:

[4]In improving paths the players are implicitly assumed to care only about the immediate benefit of deviating to the next network but they do not forecast how others might react to their actions.

Definition 3.4. An *improving path* from $h_0 \in \mathcal{H}$ to $h_k \in \mathcal{H}$ is a sequence of networks $(h_0, h_1, \ldots, h_k)$ such that for all $0 \leq l < k$ there is exactly one $c_l \in M$ with $h_{l+1} = h_l \pm (c_l, D_l)$ for some $D_l \in \mathcal{A}_{h_l}^{c_l}(\mathcal{C})$.

If there exists an improving path from $h \in \mathcal{H}$ to $h' \in \mathcal{H}$, we write $h \mapsto h'$. Moreover, let $\mathcal{I}(h) = \{h' \in \mathcal{H} \mid h \mapsto h'\}$ be the set of networks that can be reached by an improving path starting at h. Notice that h is constitutionally stable if and only if $\mathcal{I}(h) = \varnothing$. A set of networks $H \subseteq \mathcal{H}$ is *closed* if there is no improving path leading out of it, i.e., $\mathcal{I}(h) \subseteq H$ for all $h \in H$. Moreover, a set of networks $H \subseteq \mathcal{H}$ with $|H| \geq 2$ is a *cycle* if for any pair $h, h' \in H$, there exists an improving path from h to h' and vice versa.

Lemma 3.1. Let the society $(N, M, \succeq, \mathcal{C})$ be given. There exists no closed cycle if and only if, for each network $h \in \mathcal{H}$ that is not constitutionally stable, there exists an improving path leading from this network to a constitutionally stable one.

Proof. We will show the reverse statement of Lemma 3.1. If there exists a closed cycle H, by definition there exists no improving path from any $h \in H$ to a constitutionally stable network. This already proves the first direction. Now suppose there exists a network $h \in \mathcal{H}$ such that there is no constitutionally stable network in $\mathcal{I}(h)$. Therefore, this set must contain at least one cycle H_1. Suppose H_1 is a maximal cycle, i.e., it is not a proper subset of any other cycle. Now, either H_1 is closed and we are done, or it has an improving path going out of it, leading to a new maximal cycle H_2. Note that $H_1 \cap H_2 = \varnothing$. If H_2 is not closed, one can iterate the previous steps and because $\mathcal{I}(h)$ is finite, we will finally reach a closed cycle. □

Lemma 3.1 is a modification of Lemma 1 from Jackson and Watts (2002).[5] The non-existence of closed cycles not only implies existence of stable networks but it also guarantees stability in case the agents follow a "trial-and-error" strategy and care only about immediate benefits. In order to make the latter point more specific consider the following random process which has been introduced for marriage problems by Roth and Vande Vate (1990): Start with an arbitrary network $h_0 \in \mathcal{H}$. Each

This approach relates to myopic learning (e.g., Kandori *et al.*, 1993; Kandori and Rob, 1995; Monderer and Shapley, 1996) and is appropriate in relatively complex settings where it is difficult to anticipate all possible changes. In the context of coalition or network formation some authors have relaxed this assumption by analyzing "farsighted stability" (see, e.g., Page and Wooders, 2009; Page *et al.*, 2005). Conducting similar studies in our framework is left for future work.

[5] The authors have shown in slightly different terms that it is possible to find "pairwise-stable" networks if there exist no closed cycles.

round $r \in \mathbb{N}_{\geq 0}$ a pair $(c_r, D_r) \in M \times 2^N$ is drawn randomly with positive probability. If $D_r \in \mathcal{A}_{h_r}^{c_r}(\mathcal{C})$, the process moves to $h_{r+1} := h_r \pm (c_r, D_r)$. Otherwise it remains at $h_{r+1} := h_r$.

Proposition 3.1. Given a society $(N, M, \succeq, \mathcal{C})$, the random process described above always (i.e., for all $h_0 \in \mathcal{H}$) converges with probability one to a constitutionally stable network if and only if there are no closed cycles.

In the context of one-to-one matching problems, the previous result has been established by Roth and Vande Vate (1990). Although in our model the reasoning is the same, for sake of completeness we add the proof to the appendix. The intuition is straightforward. As every feasible deviation is drawn with positive probability, also every improving path has a positive probability. Therefore, if for every starting point there is an improving path leading to a constitutionally stable network, the random process converges to one of these for sure whenever it is not stopped after finitely many steps. This is particularly remarkable as in our model, network formation is not guided by a social planner or the like. Given the random process introduced above, non-existence of closed cycles is sufficient for guaranteeing that a society induces a constitutionally stable network with probability one even if the players act myopically and the deviations are not organized in a centralized way.

Proposition 3.2. Let N, M, and $\succeq$ be given. Let $\mathcal{C} \subseteq \bar{\mathcal{C}}$. Then, non-existence of closed cycles under $\bar{\mathcal{C}}$ does not imply that there are no closed cycles under $\mathcal{C}$.

Proof. In order to prove the proposition, it is sufficient to construct a suitable example. The one we consider here is a variation of an example introduced in Bogomolnaia and Jackson (2002) and revived in Diamantoudi *et al.* (2004). There are three players $N = \{i_1.i_2, i_3\}$ and one connection $M = \{c\}$. Thus, $|\mathcal{H}| = 8$. The networks are given by

	$h_1(c)$	$h_2(c)$	$h_3(c)$	$h_4(c)$	$h_5(c)$	$h_6(c)$	$h_7(c)$	$h^\varnothing$
c	$\{i_1\}$	$\{i_2\}$	$\{i_3\}$	$\{i_1, i_2\}$	$\{i_1, i_3\}$	$\{i_2, i_3\}$	$\{i_1, i_2, i_3\}$	$\varnothing$

and the players' preferences are

$$h_4 \succ^{i_1} h_7 \succ^{i_1} h_5 \succ^{i_1} h_1 \succ^{i_1} h_2 \sim^{i_1} h_3 \sim^{i_1} h_6 \sim^{i_1} h^\varnothing$$
$$h_6 \succ^{i_2} h_7 \succ^{i_2} h_4 \succ^{i_2} h_2 \succ^{i_2} h_1 \sim^{i_2} h_3 \sim^{i_2} h_5 \sim^{i_2} h^\varnothing$$
$$h_5 \succ^{i_3} h_7 \succ^{i_3} h_6 \succ^{i_3} h_3 \succ^{i_1} h_1 \sim^{i_3} h_2 \sim^{i_3} h_4 \sim^{i_3} h^\varnothing$$

The setting is actually not completely the same as in Bogomolnaia and Jackson (2002), because in their paper the authors study coalition formation (i.e., the set of players is always decomposed into a partition) while we have just one connection containing some of the players. However, "core stability" in their setting corresponds to constitutional stability with respect to the following constitutions $\mathcal{C} = (\mathcal{D}, \mathcal{S})$:

$$\mathcal{D}_h^c = 2^N \setminus \{\varnothing\} \quad \text{and} \quad \mathcal{S}_h^c(D) = \{S \subseteq h(c) \,|\, (h(c) \setminus D) \subseteq S,\, S \neq \varnothing\} \tag{3.2}$$

for all $c \in M$ and $h \neq h^{\varnothing}$. Given $\mathcal{C}$, a priori all modifications of the connection are feasible and a deviation $D \neq h(c)$ takes place if and only if all members of the resulting network benefit from deviating, i.e, $h \pm (c, D) \succ^i h$ for all $i \in h(c) \pm D$. This implies that players who are undesired can be dismissed if the other members agree on this. For the (pathological) special case of $D = h(c)$, it is required that at least one player has to approve the deviation in order to avoid exogenous instabilities. Now, given the constitutions as defined in (3.2), Diamantoudi *et al.* (2004) already pointed out that h_7 is the unique constitutionally stable (or "core stable", respectively) network and $H := \{h_4, h_5, h_6\}$ forms a closed cycle. In fact, once H is reached, there is no improving path leading to h_7 because the players act too myopically. However, consider the following constitutions $\bar{\mathcal{C}} = (\bar{\mathcal{D}}, \bar{\mathcal{S}})$: let

$$\bar{\mathcal{D}}_h^c = 2^N \setminus \{\varnothing\} \quad \text{and} \quad \bar{\mathcal{S}}_h^c(D) = \begin{cases} \{S \subseteq h(c) \,|\, (h(c) \setminus D) \subseteq S,\, S \neq \varnothing\}, & \text{if } D \cap h(c) \neq \varnothing \\ \{S \subseteq h(c) \,|\, S \neq \varnothing\}, & \text{if } D \cap h(c) = \varnothing \end{cases}$$

for all $c \in M$ and $h \neq h^{\varnothing}$. Here, granting access to c just needs the support of only one member of the connection. This obviously implies $\mathcal{C} \subsetneq \bar{\mathcal{C}}$ and, thus, the players have less blocking power (but note that the sets of stable networks coincide). However, in this case H, does not form a closed cycle any more because for all $h \in H$ there is always one member of c who supports deviating from h to h_7. Therefore, given $\bar{\mathcal{C}}$, there exist no closed cycles. □

Proposition 3.2 dissents Remark 3.2 in a way. In fact, concluding that more blocking power leads to more stability is too simplistic. Even if the set of constitutionally stable networks becomes larger, it could happen that all improving paths leading to them are severed and closed cycles occur.

Consequently, instead of enhancing the blocking power of the players, it is necessary to find alternative approaches for assuring that the society always induces a constitutionally stable network. To this end, consider once again the example constructed in the proof of Proposition 3.2. Examining it in detail yields that under $\bar{\mathcal{C}}$ we have

$h \mapsto h_7$ for all networks $h \neq h_7$ but $\mathcal{I}(h_7) = \varnothing$. Therefore, for all $h \in \mathcal{H}$ there exists a unique element in $\mathcal{I}(h)$ which is maximal with respect to "$\mapsto$". On the other hand, this is not true under $\mathcal{C}$ because $H = \{h_4, h_5, h_6\}$ forms a closed cycle and, thus, $\mathcal{I}(h) = H$ for all $h \in H$. Although these observations are limited to this specific example, similar considerations also apply in general.

Definition 3.5. Given a society $(N, M, \succeq, \mathcal{C})$, a *common ranking* $\trianglerighteq$ is a complete and transitive ordering over $\mathcal{H}$ such that $D \in \mathcal{A}_h^c(\mathcal{C})$ implies $h \pm (c, D) \trianglerighteq h$ for all $h \in \mathcal{H}$ and $c \in M$.

A common ranking $\trianglerighteq$ reflects a certain level of consensus between the players. The main idea is that the set of networks can be decomposed into several equivalence classes and once a higher class is reached, this will not be reversed afterwards. Indeed, a deviation takes place only if the joining and supporting players agree that the resulting network is not contained in a lower class than the current one. Note that a priori this is not a restriction at all because it would be possible, for instance, to choose $\trianglerighteq$ in such a way that all networks are equivalent (i.e., $h \trianglerighteq h'$ as well as $h' \trianglerighteq h$ for all $h, h' \in \mathcal{H}$). This immediately implies that a (not necessarily unique) common ranking always exists. However, the more consensus about beneficial deviations between the players, the stronger the restrictions that can be imposed by a common ranking.

Proposition 3.3. Let a society $(N, M, \succeq, \mathcal{C})$ be given.

(i) There are no cycles if and only if there exists a common ranking $\trianglerighteq$ such that for all $H \subseteq \mathcal{H}$ there is a unique $\trianglerighteq$–maximal network $\hat{h} \in H$.

(ii) There are no closed cycles if and only if there exists a common ranking $\trianglerighteq$ such that for all closed $H \subseteq \mathcal{H}$ there is a unique $\trianglerighteq$–maximal network $\hat{h} \in H$.

For the proof refer to the appendix. The main advantage of Proposition 3.3 is that it provides an alternative criterion for guaranteeing convergence to a constitutionally stable network. Item (i) states that requiring the absence of cycles is equivalent to requiring the existence of a special common ranking which identifies a unique maximal element in every subset of networks.[6] Moreover, according to (ii), having

[6] A common ranking meets this requirement if and only if it is strict. In this case, it is a variation of "Generalized Ordinal Potentials" introduced by Monderer and Shapley (1996). In particular, item (i) of Proposition 3.3 is closely related to Lemma 2.5 from their publication. Moreover, it also relates to Theorem 1 in Jackson and Watts (2001).

this feature only in particular subsets of $\mathcal{H}$ is still strong enough for excluding closed cycles. Therefore, the society induces a constitutionally stable network for sure if and only if the constitutions allow for a common ranking which is sufficiently restrictive. That is, there must be a certain degree of consent about which feasible deviations are beneficial and which are not.

3.3.1 Constitutional Rules and Players' Preferences

The remainder of this section is devoted to requirements assuring the existence of a common ranking which excludes closed cycles. In order to get more intuition for this, let us consider a stylized example:

Example 3.3. *Suppose there are three players $N = \{i_1, i_2, i_3\}$ and a unique connection $M = \{c\}$. Analogously to the example in the proof of Proposition 3.2 let $h_3(c) = \{i_3\}$, $h_5(c) = \{i_1, i_3\}$, $h_6(c) = \{i_2, i_3\}$, and $h_7(c) = \{i_1, i_2, i_3\}$. But here, the corresponding feasible deviations are $D^c_{h_3} = D^c_{h_5} = D^c_{h_6} = D^c_{h_7} = \{\{i_1\}, \{i_2\}, \{i_3\}\}$, while the supporting coalitions are given by $\mathcal{S}^c_{h_l}(D) = \{S \subseteq h_l(c) \,|\, i_3 \in S\}$ for all $D \in \mathcal{D}^c_{h_l}$ where $l \in \{3, 5, 7\}$ and $\mathcal{S}^c_{h_6}(D) = \{S \subseteq h_6(c) \,|\, i_2 \in S\}$ for all $D \in \mathcal{D}^c_{h_6}$. Moreover, the players' preferences are supposed to be as follows:*

$$h_7 \succ^{i_1} h_5 \succ^{i_1} h_6 \sim^{i_1} h_3 \succ^{i_1} \dots$$
$$h_7 \succ^{i_2} h_6 \succ^{i_2} h_5 \sim^{i_2} h_3 \succ^{i_2} \dots$$
$$h_6 \succ^{i_3} h_3 \succ^{i_3} h_5 \succ^{i_3} h_7 \succ^{i_3} \dots$$

It is not difficult to check that in this case the set $H = \{h_6, h_3, h_5, h_3\}$ forms a closed cycle because $(h_3, h_6, h_7, h_5, h_3)$ is an improving path (see Figure 3.2).
Inspecting this cycle in detail we can find a kind of irregularity in the constitutions: In h_3, h_5, and h_7, player i_3 is the only one who may decide about deviations and she even has the power to exclude the other players from the connection. But after moving to h_6, player i_3 looses her strong property rights and i_2 is able to grant i_1 access to the connection. Moreover, not only the constitutions exhibit a kind of irregularity but the players also disagree about the optimal form of the connection. First, as mentioned before, i_3 can exclude i_1 or i_2 in h_7 against their will. If either this was not possible or the players agreed to being excluded and did not want to join the connection again, the cycle would be splintered. Second, both players, i_2 and i_3, have the power to support a deviation of player i_1. The salient point is that both disagree about whether i_1 should be a member of the connection or not. If there would be a common agreement about this, one of the deviations would be blocked.

$$
\begin{tikzcd}
h_3 \arrow[r, "h_3 + (c, \{i_2\})"] & h_6 \arrow[d, bend left, "h_6 + (c, \{i_1\})"] \\
h_5 \arrow[u, "h_5 - (c, \{i_1\})"] & h_7 \arrow[l, "h_7 - (c, \{i_2\})"] \arrow[u, bend left, "h_7 - (c, \{i_1\})"]
\end{tikzcd}
$$

Figure 3.2: The cycle H

As the example illustrates, in general there are three main factors which support the occurrence of closed cycles:

(i) constitutions might change strongly even if the network itself does not,

(ii) players might be forced to leave a connection against their will, and

(iii) there might be disagreement between the players who may decide about the deviations.

In fact, for guaranteeing the existence of a common ranking which satisfies the criterion formalized in Proposition 3.3(ii), it is necessary to control for all of these factors. This implies that we not only have to find reasonable restrictions on players' preferences but also consistency conditions on the constitutions.

Definition 3.6. Given a closed set $H \subseteq \mathcal{H}$, the constitutions $\mathcal{C} = (\mathcal{D}, \mathcal{S})$ satisfy ...

- ... *regularity* with respect to H if for all $h \in H$ and $c \in M$ we have:
  (i) If $\bar{h}(c) = h(c) \cup \bar{D}$ for some $\bar{h} \in H$ and $\bar{D} \subseteq N \setminus h(c)$, then $\mathcal{D}_h^c = \mathcal{D}_{\bar{h}}^c$ and for all $D \in \mathcal{D}_{\bar{h}}^c$ and $\bar{S} \in \mathcal{S}_{\bar{h}}^c(D)$ there exists $S \in \mathcal{S}_h^c(D)$ with $S \subseteq \bar{S} \subseteq S \cup \bar{D}$.
  (ii) If $D \in \mathcal{D}_h^c$ and $S \in \mathcal{S}_h^c(D)$ with $S \notin D$, then $h(c) \setminus (S \cup D) \notin \mathcal{S}_h^c(D)$.
- ... *protection against eviction* with respect to H if for all $h \in H$ and $c \in M$ always $D \cap h(c) \subseteq S$ for all $D \in \mathcal{D}_h^c$ and $S \in \mathcal{S}_h^c(D)$.
- ... *decomposability* with respect to H if for all $h \in H$ and $c \in M$, we have that $D \in \mathcal{D}_h^c$ implies $D' \in \mathcal{D}_h^c$ and $\mathcal{S}_h^c(D) = \mathcal{S}_h^c(D')$ for all $D' \subseteq D$.

The main motivation of *regularity* is to exclude the possibility of skipping back and forth between two networks the whole time: Condition (i) states that the feasible deviations and corresponding supporting coalitions of each $c \in M$ may not vary extremely whenever c changes. If further players are added to the connection, the feasible deviations are supposed to remain the same and supporting coalitions change only as long as they might be complemented by new members. Thus, together with (ii) this implies that if a coalition $S \in \mathcal{S}_h^c(D)$ has the authority to support a deviation $D \in \mathcal{D}_h^c$, this cannot be reversed by another coalition which is neither associated to S nor to D.
If the constitutions satisfy *protection against eviction*, no player can be forced to leave a connection $c \in M$ if she does not want to do it. Modifying c always requires the consent of all deviating players (not only the consent of players who join the connection).
Decomposable constitutions exhibit a kind of independence property. If the deviation of a group of players is feasible, deviations of any subgroup of players are feasible as well and the corresponding supporting coalitions do not change.

Definition 3.7. A preference profile $\succeq$ ...

- ... satisfies *self-concern* if $h \sim^i \bar{h}$ for all $i \in N$ and each pair of networks $h, \bar{h} \in \mathcal{H}$ with $M_h(i) = M_{\bar{h}}(i)$ and $h(c) = \bar{h}(c)$ for all $c \in M_h(i)$.
- ... is *lexicographic* if each agent $i \in N$ has a preference ordering $\bar{\succeq}^i$ over 2^M such that $M_h(i) \,\bar{\succeq}^i\, M_{\bar{h}}(i)$ implies $h \succeq^i \bar{h}$ for all $h, \bar{h} \in \mathcal{H}$ with $M_h(i) \neq M_{\bar{h}}(i)$.
- ... is *uniform* if for all $i \in N$, $c \in M$, and $h, \bar{h} \in \mathcal{H}$ with $i \in h(c) = \bar{h}(c)$, $h - (c, \{k\}) \succ^j h$ implies $\bar{h} - (c, \{k\}) \succ^i \bar{h}$ and $h \succ^j h - (c, \{k\})$ implies $\bar{h} \succ^i \bar{h} - (c, \{k\})$ for $j \in h(c)$, $k \in h(c) \setminus \{i, j\}$.
- ... is *equable* if for all $i \in N$, $c \in M$, and $h, \bar{h} \in \mathcal{H}$ with $i \in h(c) = \bar{h}(c)$, $h \succ^j h - (c, \{j\})$ for some $j \in h(c)$ implies $\bar{h} \succ^i \bar{h} - (c, \{i\})$ and $h - (c, \{j\}) \succ^j h$ for some $j \in h(c)$ implies $\bar{h} - (c, \{i\}) \succ^i \bar{h}$.
- ... is *separable* if for all $i \in N$, $c \in M$, and $h, \bar{h} \in \mathcal{H}$ with $i \in h(c) \subseteq \bar{h}(c)$ the two following conditions are satisfied:
  (i) $\bar{h} - (c, D) \succ^i \bar{h}$ if and only if $h - (c, D) \succ^i h$ for all $\emptyset \neq D \subseteq h(c) \setminus \{i\}$.
  (ii) $\bar{h} + (c, D) \succ^i \bar{h}$ if and only if $h + (c, D) \succ^i h$ for all $\emptyset \neq D \subseteq N \setminus \bar{h}(c)$.

Self-concern is a kind of independence property. Player i neither benefits nor suffers if the network changes in such a way that i is not affected directly.
The definition of *lexicographic* preferences is adapted from Dutta and Massó (1997). Under this requirement, each player $i \in N$ is mainly concerned about the connections themselves where she is a member of and less about who the other members are. Only if $M_h(i) = M_{\bar{h}}(i)$, might she care about the other players in her connections.
If the preferences of the players are *uniform* and a player leaves a connection, either all remaining members benefit from this deviation or none of them. Note that this is supposed to be independent of the form the other connections have.
Under *equability* player $i \in N$ wants to stay in a connection $c \in M$ only if the other members also want to stay. Suppose, for example, the connections generate a payoff which is distributed equally among the members: Then, if a player has an incentive to leave c, the same goes for i.
Separability as introduced here is a variation of the same-named concept from Banerjee *et al.* (2001). The idea is that player i's support for a certain leaving or joining group D is independent of the form the connection actually has.

3.3.2 Non-existence of (Closed) Cycles

Now, combining the restrictions introduced in the previous subsection allows formulating conditions which guarantee non-existence of (closed) cycles and thus lead to convergence to a constitutionally stable network.

Proposition 3.4. Let a society $(N, M, \succeq, \mathcal{C})$ be given where all constitutions satisfy protection against eviction with respect to a closed set $H \subseteq \mathcal{H}$. If the players' preferences satisfy equability and self-concern, there exist no cycles in H.

All proofs of this subsection can be found in the appendix. The requirements of Proposition 3.4 reflect the three factors which might cause instabilities. Equability and self-concern, for example, impose restrictions on the players' preferences. Both conditions together guarantee that there is only little disagreement about the optimal form of each connection $c \in M$. Complementing this, protection against eviction with respect to H has two consequences: On the one hand, as the definition directly implies, players cannot be forced to leave a connection if they do not agree to this. On the other hand, indirectly it also ensures that the constitutions do not change too strongly whenever a connection is altered. More specifically, $S \in \mathcal{S}_h^c(D)$ implies $h(c) \setminus S \notin \mathcal{S}_h^c(D)$ for all $h \in H$, $c \in M$, and $D \in \mathcal{D}_h^c$. The interpretation is similar to regularity: If S has the power to support a deviation of D, this cannot be reversed by other supporting coalitions.

Proposition 3.5. Let a society $(N, M, \succeq, \mathcal{C})$ be given where all constitutions satisfy protection against eviction with respect to a closed set $H \subseteq \mathcal{H}$. If the players' preferences are lexicographic, there exist no cycles in H.

The intuition of the previous result is similar to Proposition 3.4. Obviously, the only difference is that the preferences are not supposed to satisfy equability and self-concern but lexicography instead. Therefore, even if there is some disagreement about the optimal form of the connections, it is relegated to a secondary role.
Both previous propositions exclude the existence of not only closed cycles but even of cycles in general. To some extent this is caused by protection against eviction. Indeed, it is not possible to drop or to relax this assumption without reinforcing the requirements on players' preferences.

Proposition 3.6. Let a society $(N, M, \succeq, \mathcal{C})$ be given. Assume all constitutions are decomposable and regular with respect to a closed set $H \subseteq \mathcal{H}$. Moreover, suppose the players' preferences are separable, uniform, equable, and they satisfy self-concern. Then, there exist no closed cycles in H.

As the definition directly implies, regularity inhibits the constitutions from varying too extremely and, similar to Proposition 3.4, equability and self-concern guarantee a certain degree of consent about the optimal form of the network. In addition to this, due to separability and uniformity, in most networks the players are not forced to leave their connections if they do not agree to this. If, for example, some player's entry is supported by a certain coalition, the corresponding members will not change their minds, even if the connection is altered strongly. Thus, the player will only leave again if she has an incentive for deviating.
Note that similar to Proposition 3.4, it is required that the preferences satisfy equability and self-concern together. Consequently, and analogously to above, it is possible to replace both assumptions in Proposition 3.6 by lexicography. The intuition is the same: The optimal form of the connections is relegated to a secondary role.

Proposition 3.7. Let the society $(N, M, \succeq, \mathcal{C})$ be given. Assume all constitutions are decomposable and regular with respect to a closed set $H \subseteq \mathcal{H}$. Moreover, suppose the preferences of the players are separable, uniform and lexicographic. Then, there exist no closed cycles in H.[7]

[7] Although the proof proceeds similarly as the one of Proposition 3.6, the main idea is partially

3.4 Many-to-many Matching Markets

One of the most intriguing features of our model is a versatile applicability since overlapping group structures are characteristic for many environments. Consider, for example, many-to-many matching markets. The main primitives of these markets are two finite sets of players E and F, where the members of E are usually interpreted as employees (or workers) and the members of F as firms (see, e.g., Roth, 1984). A (two-sided) many-to-many matching $\mu \subseteq E \times F$ is then simply a collection of worker-firm pairs indicating which employees are working for which firms. Both sides of the market, i.e., all players in E as well as all players in F, are supposed to have preferences over all possible matchings. Thereby, the employees are classically assumed to care only about which firms they work for but not about who their co-workers might be. The owners, on the other hand, are only concerned about the employees working for their firm:

> "This involves an assumption that workers are indifferent to who their co-workers might be, and firms are indifferent to whether their employees moonlight at other jobs." (Roth, 1984, p. 51)

The setting outlined above can be embedded into our model in a straightforward way: Let $M := F$, i.e., each connection $c \in M$ is interpreted as firm. Since in our model the connections do not act as players, we suppose that each firm c has exactly one owner $o_c \in O$. That is, in the following we assume that the set of players $N := E \cup O$ can be decomposed into two (disjoint) subsets, the employees E and the owners O. Given these preliminaries, each matching $\mu \subseteq E \times F$ can be represented by the social network $h^\mu \in H$ which is defined via $h^\mu(c) = \{i \in E \,|\, (i,c) \in \mu\} \cup \{o_c\}$ for all $c \in M = F$. In order to be in line with classical literature on many-to-many matchings, we assume that each owner has no incentive for leaving her firm or for joining any other firm, i.e., we are only interested in the case $O \cap h(c) = \{o_c\}$ for $h \in \mathcal{H}$ and $c \in M$.[8] Nevertheless, since a priori we do not exclude certain network structures, for technical reasons we also have to define preferences over networks where this requirement is not met. To fix ideas, Roth's assumptions on the players' preferences imply that each employee $i \in E$ is indifferent to all networks where she is working for the same set of firms, i.e., $h \sim^i \bar{h}$ for all $h, \bar{h} \in \mathcal{H}$ with $M_h(i) = M_{\bar{h}}(i)$.

based on Section 5 of Sotomayor (1999).

[8] Thus, we do not consider the possibility of changing the owner. But from a technical point of view it would not be difficult to include this feature into the model.

Moreover, given $c \in M$ and $O \cap h(c) = \{o_c\}$, the assumptions also imply $h \sim^{o_c} \bar{h}$ whenever $h(c) = \bar{h}(c)$. For the (pathological) case of $O \cap h(c) \neq \{o_c\}$, we assume $h \pm ((O \cap h(c)) \pm \{o_c\}) \succ^{o_c} h$. Therefore, the preferences of all employees are lexicographic; and restricted to the set $H := \{h \in \mathcal{H} \mid O \cap h(c) = \{o_c\} \ \forall \ c \in M\}$ the same goes for the owners, too.

Since our model is richer than the classical matching approach (in particular, social networks as defined here might be interpreted as *one-sided* many-to-many matchings), it consequently enables us to model job markets more realistically. Complementing this, our formalization of constitutions supports plausibility even further as it allows studying different levels of authority of the owners in a flexible way. For instance, in many countries (especially in Europe) employees are organized in labor unions which represent the interests of their members. These unions may guarantee a quite strong protection against dismissal to the employees and in the short run the consent of a worker is needed if the owner wants her to leave the firm. Many-to-many matching theory, however, usually concentrates on job markets without strong protection against dismissal, like the US job market, for example, and neglects the impact of labor unions. Due to its versatility, our model provides an appropriate framework for examining and comparing these settings in a convenient way. Therefore, the remainder of the paper is devoted to the study of different levels of authority of the owners.

3.4.1 Protection against Dismissal

In the following, we always assume that the employees are allowed to accept as many jobs as they want to. Moreover, the firms have unlimited capacity to hire workers, i.e., given $O \cap h(c) = \{o_c\}$ for $h \in \mathcal{H}$ and all $c \in M$, every possible deviation of the employees is feasible. Nevertheless, quotas could be included easily by allowing only for deviations which respect a maximal firm size. For sake of completeness, we also have to consider the case where an owner is not part of her firm or other owners are contained in it. Then, we assume that the only feasible deviation is to add the owner and to delete all the others:

$$\mathcal{D}_h^c = \begin{cases} 2^E, & \text{if } O \cap h(c) = \{o_c\} \\ (O \cap h(c)) \pm \{o_c\}, & \text{if } O \cap h(c) \neq \{o_c\} \end{cases} \tag{3.3}$$

As mentioned before, in this subsection we suppose that the owner has no authority to fire her employees if they do not agree with this. However, she is the only one who may decide about hiring new workers. On the other hand, each employee is free to terminate her job if she has an incentive to do so. These considerations lead to the following set of supporting coalitions:

$$\mathcal{S}_h^c(D) = \begin{cases} \{S \subseteq h(c) \mid D \cap h(c) \subseteq S \text{ and } o_c \in S\}, & \text{if } O \cap h(c) = \{o_c\} \text{ and } D \nsubseteq h(c) \\ \{S \subseteq h(c) \mid D \subseteq S\}, & \text{if } O \cap h(c) = \{o_c\} \text{ and } D \subseteq h(c) \\ \{\varnothing\}, & \text{if } O \cap h(c) \neq \{o_c\} \end{cases}$$

Note that for the case of $O \cap h(c) \neq \{o_c\}$, we assume that the empty set is the only supporting coalition and, thus, these networks are not stable by construction.

Corollary 3.1. There are no cycles in "Protection against Dismissal".

Proof. This follows immediately from Proposition 3.5 because the players' preferences are lexicographic and we also have protection against dismissal with respect to the closed set H. □

At first sight, the previous result might be slightly surprising because in many studies about two-sided many-to-many matchings the existence of stable structures is an issue (e.g., Sotomayor, 2004). This is mainly due to the fact that this literature normally examines environments where the owners are free to fire a worker if they benefit from it. Indeed, protection against dismissal is the driving force of the previous result. Let $\mathcal{ST}^{\text{PD}}$ denote the set of stable networks in Protection against Dismissal. Note that this set also contains the *worker-optimal networks* which are defined as follows: Suppose $\bar{M}^i \subseteq M$ is a set of firms which is mostly preferred by $i \in E$. Then, if h^{wo} is given by $h^{wo}(c) = \{i \in E \mid c \in \bar{M}^i\} \cup \{o_c\}$ for all $c \in M$, every employee is assigned to a set of firms she preferably wants to work for and, thus, she obviously has no incentive for deviating.

3.4.2 Hire and Fire

Let us now focus on job markets without strong protection against dismissal. Translated to the model considered here, this means that the owners have the right to fire employees even if these do not agree to leaving. This aspect can be captured by considering the following supporting coalitions:

$$\mathcal{S}_h^c(D) = \begin{cases} \{S \subseteq h(c) \mid o_c \in S\}, & \text{if } O \cap h(c) = \{o_c\}, D \nsubseteq h(c) \\ \{S \subseteq h(c) \mid D \subseteq S \text{ or } o_c \in S\}, & \text{if } O \cap h(c) = \{o_c\}, D \subseteq h(c) \\ \{\varnothing\}, & \text{if } O \cap h(c) \neq \{o_c\} \end{cases}$$

Let $\mathcal{ST}^{\mathrm{HF}}$ be the set of stable networks in "Hire and Fire". Note that Remark 3.2 implies $\mathcal{ST}^{\mathrm{HF}} \subseteq \mathcal{ST}^{\mathrm{PD}}$. However, it is well known that without further assumptions existence of stable networks in Hire and Fire is not assured (as can be easily seen by means of an example with two workers and two firms). Thus, in order to exclude existence of closed cycles it is necessary to impose further restrictions on constitutions or preferences. A straightforward approach would be to proceed similarly to Proposition 3.7 since, as mentioned above, the preferences of the players are lexicographic and, moreover, the constitutions in Hire and Fire are not only decomposable but also regular with respect to H. However, due to the specific structure of the setting considered here it is not necessary to impose such strong assumptions as in Proposition 3.7. Since the owners are the only players who have the authority to hire new employees and because they never leave their firm, uniformity is not needed and it is sufficient to additionally assume that the owners' preferences are separable.

Proposition 3.8. If the preferences of the owners are separable, there exist no closed cycles in Hire and Fire.

Remark 3.3. This proposition is in line with several other well-known publications from literature, like the papers from Roth and Vande Vate (1990), Chung (2000), Diamantoudi *et al.* (2004), and especially Kojima and Ünver (2008). Similar to our result, Kojima and Ünver (2008) have shown in the context of two-sided many-to-many matchings that if employees and owners have, respectively, "substitutable" (see Roth, 1984) and "responsive" (see Roth, 1985) preferences, there always exists an improving path leading to a "pairwise stable" matching. In fact, the assumptions we impose in Proposition 3.8 are less restrictive: Given preferences as defined at the beginning of this section and if, in addition to this, only deviations of single players are feasible, responsiveness of the owners' preferences implies separability which in turn implies substitutability (converse implications do not hold). Therefore, Proposition 3.8 complements their findings.

Although we have $\mathcal{ST}^{\mathrm{HF}} \subseteq \mathcal{ST}^{\mathrm{PD}}$, the converse inclusion does not necessarily hold. Therefore, there might exist networks which are stable in Protection against Dismissal but would be blocked if the owners' level of authority is sufficiently high. In particular, due to the characteristics of Hire and Fire, if $h \in \mathcal{ST}^{\mathrm{PD}}$ but $h \notin \mathcal{ST}^{\mathrm{HF}}$, there is at least one owner who would like to fire some of her employees. This already indicates that the interests of both sides of the market might be opposed in some way. For deepening these considerations further we need to enhance separability:

Definition 3.8. A preference profile $\succeq$ is strongly separable if for all $i \in N$, $c \in M$, and $h, \bar{h} \in \mathcal{H}$ with $i \in h(c) \subseteq \bar{h}(c)$, the two following conditions are satisfied:

(i) $\bar{h} - (c, D) \succ^i \bar{h}$ if and only if $h - (c, D) \succ^i h$ for all $\emptyset \neq D \subseteq h(c)$.

(ii) $\bar{h} + (c, D) \succ^i \bar{h}$ if and only if $h + (c, D) \succ^i h$ for all $\emptyset \neq D \subseteq N \setminus \bar{h}(c)$.

As the name implies, strong separability is a stronger requirement than separability. Again, player i's support for a certain leaving or joining group is independent of the form the connection actually has. But under strong separability this is also true if i belongs to the deviating group, i.e., if i leaves the connection. Translated to Hire and Fire, this basically means that i's preference about whether to work for a firm $c \in M$ or not is independent of the other firms she is working for.

Proposition 3.9. Assume the workers' preferences are strongly separable and the owners' preferences are separable. Moreover, suppose the worker-optimal network h^{wo} is uniquely determined. Then, $h^{wo} \in \mathcal{ST}^{\text{HF}}$ if and only if $\mathcal{ST}^{\text{PD}} = \mathcal{ST}^{\text{HF}}$.

Proof. If $\mathcal{ST}^{\text{PD}} = \mathcal{ST}^{\text{HF}}$, then also $h^{wo} \in \mathcal{ST}^{\text{HF}}$ because h^{wo} is always stable in Protection against Dismissal and there remains nothing to show. For the other direction, suppose the statement is not true, i.e., $h^{wo} \in \mathcal{ST}^{\text{HF}}$ but $\mathcal{ST}^{\text{HF}} \subsetneq \mathcal{ST}^{\text{PD}}$. Let $\bar{h} \in \mathcal{ST}^{\text{PD}} \setminus \mathcal{ST}^{\text{HF}}$. Then, there must be an owner o_c who would block $\bar{h}$ if her property rights are strong enough, i.e., there exists an employee $i \in \bar{h}(c)$ such that $\bar{h} - (c, \{i\}) \succ^{o_c} \bar{h}$. Because o_c's preferences are separable and h^{wo} is stable, this implies $i \notin h^{wo}(c)$. Otherwise, the owner would also have an incentive to dismiss the employee in h^{wo}. Thus, uniqueness of h^{wo} yields that i strictly prefers h^{wo} to $h^{wo} + (c, \{i\})$. In particular, because her preferences are supposed to be separable, she would also have a strict incentive for leaving c at $\bar{h}$, but this contradicts stability of this network. $\square$

Proposition 3.9 is in line with Roth (1984). Under the requirement that the preferences of owners and employees are "substitutable", the author finds a "conflict of interest between agents on opposite sides [of the market]" (Roth (1984), p. 47). A similar conflict also arises here: Given (strong) separability of the players' preferences, the stable outcome which would be blocked first by the owners is the worker-optimal network. In fact, this phenomenon is completely independent of specific working conditions such as wages or the working environment, for example, because we abstracted from factors like these. Moreover, as will be shown in the following, the conflict becomes even stronger if the owners' level of authority is raised higher.

3.4.3 Slavery

Roughly speaking, "Slavery" is the counterposition of Protection against Dismissal. Here, the owners not only have the power to decide about new employees but also about whether workers may leave their firm or not:

$$\mathcal{S}_h^c(D) = \begin{cases} \{S \subseteq h(c) \mid o_c \in S\}, & \text{if } O \cap h(c) = \{o_c\} \text{ and } D \nsubseteq h(c) \\ \{S \subseteq h(c) \mid o_c \in S\}, & \text{if } O \cap h(c) = \{o_c\} \text{ and } D \subseteq h(c) \\ \{\varnothing\}, & \text{if } O \cap h(c) \neq \{o_c\} \end{cases}$$

By applying Proposition 3.5 we get the following result:

Corollary 3.2. Every improving path in Slavery leads to a constitutionally stable network.

Let $\mathcal{ST}^{\text{SL}}$ be the corresponding set of stable networks.

Remark 3.4. It is easy to check that a network is stable in Hire and Fire if and only if it is stable in Protection against Dismissal and Slavery, i.e., $\mathcal{ST}^{\text{HF}} = \mathcal{ST}^{\text{PD}} \cap \mathcal{ST}^{\text{SL}}$. But it might be the case that the intersection of the sets of stable networks is empty. However, according to Corollary 3.1 and Corollary 3.2 there exist no cycles in Protection against Dismissal and Slavery. Therefore, a simple algorithm for finding stable networks in Hire and Fire (in case they exist) is to determine the sets of maximal elements of all improving paths in the two other settings and to check whether the intersection of these sets is non-empty.

Analogously to worker-optimal networks it is, of course, also possible to define firm-optimal networks. Let $\hat{E}^c \subseteq E$ be a set of employees which is mostly preferred by player o_c and define h^{fo} by $h^{fo}(c) = \hat{E}^c \cup \{o_c\}$ for all $c \in M$. Then, none of the owners has an incentive for deviating and, thus, the network is stable in Slavery.

Proposition 3.10. Assume the workers' preferences are strongly separable and the owners' preferences are separable. Moreover, suppose the firm-optimal network h^{fo} is uniquely determined. Then, $h^{fo} \in \mathcal{ST}^{\text{HF}}$ if and only if $\mathcal{ST}^{\text{SL}} = \mathcal{ST}^{\text{HF}}$.

Proof. Because Slavery is symmetric to Protection against Dismissal, the proof proceeds analogously to the one of Proposition 3.9, just by reversing the role of owners and employees. □

Proposition 3.10 has two implications: First, it shows that the owners can enforce the network which is most beneficial for them if they have a high level of authority. Second, this network would be the first network which is rejected by the employees. In fact, this result extends and reinforces the interpretation of Proposition 3.9 in a straightforward way: Each side of the market would be worse off if the other side obtains more property rights. If, for example, labor unions narrow the owners' level of authority, the employees would benefit from this and vice versa. Recall once more that this insight does not directly depend on wages or the like, which we do not consider explicitly in our model. In particular, this implies that Roth's "polarization of interests" seems to achieve great generality.

3.5 Conclusion

Even though there is an immense and rich body of literature on the stability of networks (or group structures, respectively), in most of these studies, the stability concepts the authors use are relatively rigid since they do not consider explicitly the rules governing network formation. Indeed, the most distinctive feature of our framework is the formal introduction of constitutions which enable modeling these rules in a very flexible way. Using this approach we find that enhancing the blocking power of the players does not necessarily lead to more stability. Moreover, we show that the society induces a constitutionally stable network for sure if and only if there is a certain degree of consent between the players about which feasible deviations (according to the constitutions) are beneficial and which are not. In this context, we also discuss conditions under which this criterion is satisfied. By applying our model to job markets with labor unions we find a variation of Roth's "polarization of interests": The workers generically suffer if the degree of authority of the owners is raised and vice versa. In addition to this, we also show that the markets always become stable if the property rights of one side of the market become sufficiently strong.

Although the model we analyze in this paper expands well-established branches like Network Theory or Matching Theory, for example, it is still subject to certain limitations which narrow the field of possible applications. For instance, assuming myopic behavior is reasonable for a start, but it is well-justified only in complex settings where it is extremely difficult to anticipate all possible alterations. Therefore, it might be worth analyzing which results could be obtained if the players act farsightedly. Another natural extension is to examine situations where it is possible to add new players or connections to the society. To incorporate this kind of dynamics, it would be necessary to relax the assumption of fixed sets of players and connections. Furthermore, under certain requirements common rankings relate to ordinal potentials. Since there are numerous publications on potential functions (e.g., Hart and Mas-Colell, 1989; Monderer and Shapley, 1996; Page and Wooders, 2010; Qin, 1996; Slikker, 2001), it seems interesting to study whether the corresponding results also extend to the model introduced here.

Acknowledgments

We thank Nizar Allouch, Sonja Brangewitz, Herbert Dawid, Claus-Jochen Haake, Tim Hellmann, Christoph Kuzmics, and Jakob Landwehr for useful comments and suggestions. We are also thank session participants at the "5th Edition of the UECE Lisbon Meetings in Game Theory and Applications" as well as the seminar audience at the research seminar in Bielefeld. Ana Mauleon and Vincent Vannetelbosch are Senior Research Associates of the National Fund for Scientific Research (FNRS). Financial support from the Spanish Ministry of Economy and Competition under the project ECO 2012-35820 is gratefully acknowledged. Large parts of this work were done when Nils Roehl was visiting the Center for Operations Research and Econometrics (CORE) in Louvain-la-Neuve and he gratefully acknowledges hospitality. He is also grateful for financial support by the German Research Foundation (DFG) within the Collaborative Research Center "On-The-Fly Computing" (SFB 901).

3.6 Appendix: Proofs

Proof of Proposition 3.1

Basically speaking, the proposition is an immediate implication of standard results in Probability Theory. For a formal introduction see Grimmett and Stirzaker (2001), for example. For each pair $(c, D) \in M \times 2^N$ let $p_{c,D} > 0$ denote the probability of choosing (c, D). Then, given a network $h \in \mathcal{H}$, the probability $p_{hh'} \in [0, 1]$ of moving from h to another network $h' \neq h$ is

$$p_{hh'} = \begin{cases} p_{c,D}, & \text{if } h' = h \pm (c, D) \text{ for some } (c, D) \in M \times 2^N \text{ and } D \in \mathcal{A}_h^c(\mathcal{C}) \\ 0, & \text{if either } h' \neq h \pm (c, D) \text{ for all } (c, D) \in M \times 2^N \text{ or } D \notin \mathcal{A}_h^c(\mathcal{C}). \end{cases}$$

Moreover, we define $p_{hh} \in [0,1]$ via $p_{hh} = 1 - \sum_{h':h' \neq h} p_{hh'}$. Let $X = (X_r)_{r \in \mathbb{N}_{\geq 0}}$ be the (homogeneous) Markov Chain describing the random process introduced in Section 3.3, i.e, the probability of moving from $h_r \in \mathcal{H}$ to $h_{r+1} \in \mathcal{H}$ in round $r \geq 0$ is given by

$$\mathbb{P}\left(X_{r+1} = h' | X_0 = h_0, X_1 = h_1, \ldots, X_r = h_r\right) = \mathbb{P}\left(X_{r+1} = h' | X_r = h_r\right) = p_{h_r h_{r+1}}$$

for all $h_0, \ldots, h_{r-1} \in \mathcal{H}$. Therefore, $\mathbb{P}(X_{r+1} = \bar{h} | X_r = \bar{h}) = 1$ for all $r \geq 0$ if and only if $\bar{h}$ is constitutionally stable. In particular, this implies that each constitutionally stable network is persistent, where a network $h \in \mathcal{H}$ is said to be *persistent* if $\mathbb{P}\left(X_r = h \text{ for some } r \geq 1 | X_0 = h\right) = 1$ (cf. Grimmett and Stirzaker, 2001). If h is not persistent it is called *transient*. For each pair of networks $h, h' \in \mathcal{H}$, let $p_{hh'}(s) = \mathbb{P}(X_{r+s} = h' | X_r = h)$ be the probability of reaching h' from h in s steps. According to the Chapman-Kolmogorov equation, this probability does not depend on r and so is well-defined indeed.

The definitions introduced in the preceding paragraph provide a basis for proving the proposition formally. For the first direction let a closed cycle H be given. Thus, we have $p_{hh'} = 0$ for all $h \in H$ and $h' \notin H$. Let $(h_0, h_1, \ldots, h_k)$ be an arbitrary sequence of networks with $h_0 \in H$ and $h_k \in \mathcal{ST}(\mathcal{C})$. Note that we do not require that the sequence is an improving path. Nevertheless, there has to be at least one $0 \leq r < k$ with $h_r \in H$ and $h_{r+1} \notin H$ and, thus, $\mathbb{P}(X_1 = h_1, \ldots, X_k = h_k, \ldots, X_k = h_k | X_0 = h_0) = \mathbb{P}(X_1 = h_1 | X_0 = h_0) \cdot \ldots \cdot \mathbb{P}(X_{r+1} = h_{r+1} | X_r = h_r) \cdot \ldots \cdot \mathbb{P}(X_k = h_k | X_{k-1} = h_{k-1}) = 0$. Since the sequence was chosen arbitrarily, this implies $p_{h_0 h_k}(s) = 0$ for all $s \geq 0$. But from this it immediately follows that the probability of converging to a constitutionally stable network is equal to zero if the Markov Chain starts in H.

Now suppose there are no closed cycles and let $h' \notin \mathcal{ST}(\mathcal{C})$ be a non-stable network. According to Lemma 3.1 there exists an improving path $(h_0, h_1, \ldots, h_k)$ with $h_0 = h'$ leading to a constitutionally stable network h_k. Then:

$$\begin{aligned}
\mathbb{P}(X_r = h' | X_0 = h') &\leq 1 - \mathbb{P}(X_1 = h_1, \ldots, X_k = h_k, \ldots, X_r = h_k, X_{r+1} = h_k | X_0 = h') \\
&= 1 - \mathbb{P}(X_1 = h_1 | X_0 = h') \cdot \ldots \cdot \mathbb{P}(X_k = h_k | X_{k-1} = h_{k-1}) \\
&\quad \cdot \underbrace{\mathbb{P}(X_{k+1} = h_k | X_k = h_k) \cdot \ldots \cdot \mathbb{P}(X_{r+1} = h_k | X_r = h_k)}_{=1} \\
&= 1 - \underbrace{\mathbb{P}(X_1 = h_1 | X_0 = h') \cdot \ldots \cdot \mathbb{P}(X_k = h_k | X_{k-1} = h_{k-1})}_{>0} < 1
\end{aligned}$$

for all $r \geq 1$. Therefore, a network h' is transient if and only if h' is not constitutionally stable and this allows applying a well-known result in Probability Theory:

Proposition (Cf. Grimmett and Stirzaker, 2001, p. 222)**.** If h' is transient then $p_{hh'}(s) \to 0$ as $s \to \infty$ for all $h \in \mathcal{H}$.

Thus, for every starting network h, the probability of converging to a transient network is equal to zero. Or, stated equivalently: the probability that the Markov Chain introduced above converges to a constitutionally stable network (or persistent network, respectively) is always one. $\square$

Proof of Proposition 3.3(i)

In order to show that the existence of $\unrhd$ implies the non-existence of cycles, we will consider the counter-position of this statement. Therefore, assume there is a cycle $H \subseteq \mathcal{H}$. Since there exists a path from each network to every other network in H, if $\unrhd$ is a common ranking, we must have $\bar{h} \unrhd \breve{h}$ as well as $\breve{h} \unrhd \bar{h}$ for all $\bar{h}, \breve{h} \in H$. Thus, there is no unique $\unrhd$-maximal element in H.
For the other direction suppose there exist no cycles. The following algorithm proceeds in a similar way as the one in the proof of Theorem 1 in Jackson and Watts (2001). We start with the binary relation $\unrhd_1$ where $h \rhd_1 \bar{h}$ if and only if there exists an improving path from $\bar{h}$ to h. Since there is no cycle, $\unrhd_1$ is strict. Moreover, for all $h \in \mathcal{H}$, $c \in M$, and $D \in \mathcal{D}^c_h$, deviating from h to $h \pm (c, D)$ always implies $h \pm (c, D) \rhd_1 h$ by construction. However, $\unrhd_1$ is not necessarily complete. Let $\tilde{h}, \breve{h} \in \mathcal{H}$ with neither $\tilde{h} \rhd_1 \breve{h}$ nor $\breve{h} \rhd_1 \tilde{h}$. We construct $\bar{\unrhd}_1$ by adding $\tilde{h} \bar{\rhd} \breve{h}$ to $\unrhd_1$, i.e., $h \bar{\unrhd}_1 \bar{h}$ if and only if $h \unrhd_1 \bar{h}$ or $h = \tilde{h}$ and $\bar{h} = \breve{h}$. Moreover, let $\unrhd_2$ be the transitive closure of $\bar{\unrhd}_1$. We will show that $\unrhd_2$ still represents the preference profile of the players, i.e., deviating from h to $h \pm (c, D)$ always implies $h \pm (c, D) \rhd_2 h$ for all $c \in M$ and $D \in \mathcal{A}^c_h(\mathcal{C})$. Suppose this is not true, that is, suppose there exist $h' \in \mathcal{H}$, $c \in M$, $D \in \mathcal{D}^c_{h'}$, and $S \in \mathcal{S}^c_{h'}(D)$ with $h' \pm (c, D) \succ^i h'$ for all $i \in (D \setminus h'(c)) \cup S$ but $h' \unrhd_2 h' \pm (c, D)$. Thus, there exists a sequence of networks $(h_0, h_1, \dots, h_k)$ with $h_0 = h'$, $h_k = h' \pm (c, D)$ and $h_0 \bar{\unrhd}_1 h_1 \bar{\unrhd}_1 \dots \bar{\unrhd}_1 h_k$. Assume the sequence is of minimal length. This implies that $h_l = h_{l'}$ only if $l = l'$ for all $l, l' \in \{0, 1, \dots, k\}$. Suppose there exists an $l \in \{1, \dots, k\}$ with $\{h_{l-1}, h_l\} = \{\breve{h}, \tilde{h}\}$. Because $h_{l'} \neq \breve{h}, \tilde{h}$ for all $l' \notin \{l-1, l\}$ this yields

$$h_l \unrhd_1 h_{l+1} \unrhd_1 \dots \unrhd_1 h_k = h' \pm (c, D) \unrhd_1 h' = h_0 \unrhd_1 \dots \unrhd_1 h_{l-1}$$

and, thus, there exists an improving path from $\tilde{h}$ to $\breve{h}$ or vice versa. This contradicts the assumption that the two networks are not comparable under $\unrhd_1$. Therefore, there exists no $l \in \{1, \dots, k\}$ with $\{h_{l-1}, h_l\} = \{\breve{h}, \tilde{h}\}$. From this follows $h_0 \unrhd_1 h_1 \unrhd_1 \dots \bar{\unrhd}_1 h_k$

which contradicts the assumption that there is no cycle. Thus, $\trianglerighteq_2$ still represents the preferences of the players and by construction it is also transitive and strict. If it is not complete, the previous steps can be iterated. As the set of networks is finite, the iteration will stop after finitely many steps and we obtain a common ranking $\trianglerighteq$ which is strict. In particular, strictness implies that for each $H \subseteq \mathcal{H}$ there is a unique $\trianglerighteq$-maximal network $\hat{h} \in H$. $\square$

Proof of Proposition 3.3(ii)

The first direction proceeds analogously to the first direction of Part (i). Let a common ranking $\trianglerighteq$ and a set of networks $H \subseteq \mathcal{H}$ be given. If H forms a closed cycle, we have $\mathcal{I}(h) = \mathcal{I}(h') = H$ and $h \trianglerighteq h'$ as well as $h' \trianglerighteq h$ for all $h, h' \in H$. But this would contradict that there is a unique $\trianglerighteq$-maximal network in H and, thus, there cannot exist a closed cycle.
For the other direction suppose there exist no closed cycles. The first step of the construction of the common ranking proceeds in the same way as the one of Part (i). That is, we start with $\trianglerighteq_1$ where $h \trianglerighteq_1 \bar{h}$ if and only if there exists an improving path from $\bar{h}$ to h. But note that here this binary relation is not necessarily strict. Since by assumption there are no closed cycles, there exists at least one constitutionally stable network $h' \in \mathcal{H}$. If this network is uniquely determined, according to Lemma 3.1 it is contained in every closed subset $H \subseteq \mathcal{H}$ and $\trianglerighteq_1$ can then obviously be extended to a complete ranking where h' is the unique maximal element. Therefore, in the following, suppose there exists a further constitutionally stable network $h'' \in \mathcal{H}$. In particular, this implies that neither $h' \trianglerighteq_1 h''$ nor $h'' \trianglerighteq_1 h'$. Let $\tilde{h}, \breve{h} \in \mathcal{H}$ be an arbitrary pair of networks not comparable under $\trianglerighteq_1$. Analogously to above, $\bar{\trianglerighteq}_1$ is constructed by adding $\tilde{h} \,\bar{\triangleright}_1\, \breve{h}$ to $\trianglerighteq_1$, i.e., $h \,\bar{\trianglerighteq}_1\, \bar{h}$ if and only if $h \trianglerighteq_1 \bar{h}$ or $h = \tilde{h}$ and $\bar{h} = \breve{h}$. Again, let $\trianglerighteq_2$ be the transitive closure of $\bar{\trianglerighteq}_1$. Note that by construction $h' \trianglerighteq_2 h''$ would imply $h' \ntrianglelefteq_2 h''$ and vice versa. If $\trianglerighteq_2$ is not complete, because of finiteness of $\mathcal{H}$ we can iterate the previous steps until a complete ranking $\trianglerighteq$ is reached. We will show that h' and h'' are still not equivalent under $\trianglerighteq$. This, in fact, has the following implication: If $\breve{h}$ is $\trianglerighteq$–maximal in a closed subset $H \subseteq \mathcal{H}$, it has to be constitutionally stable by construction and w.l.o.g. we may assume $\breve{h} = h'$. Then, for any other stable network $h'' \in H$, we must have $h' \triangleright h''$ and, thus, h' is the unique $\trianglerighteq$–maximal element in H.
In order to show that h' and h'' are still not equivalent under $\trianglerighteq$, let $\trianglerighteq_k$ be the binary relation constructed in the k-th step of the algorithm described in the previous passage. For $k = 1, 2$ we already know that $h' \trianglerighteq_k h''$ would imply $h' \ntrianglelefteq_k h''$ and vice

versa. We will show inductively that this is also satisfied for all other k. Therefore, let $k \geq 3$ and suppose that h' and h'' are not equivalent under $\trianglerighteq_{k-1}$. Moreover, assume this is not satisfied under $\trianglerighteq_k$, i.e, we have $h' \trianglerighteq_k h''$ as well as $h'' \trianglerighteq_k h'$. This assumption will lead to a contradiction. Let $\tilde{h}^{(k-1)}, \breve{h}^{(k-1)} \in \mathcal{H}$ be the corresponding pair of networks not comparable under $\trianglerighteq_{k-1}$ which is added in the next step. We will distinguish three cases:

Case 1: $h' \triangleright_{k-1} h''$.

Because we assume that h' and h'' are not equivalent under $\trianglerighteq_{k-1}$, this implies that there exists a sequence of networks $(h_1, \dots, h_l)$ with $h_1 = h''$, $h_l = h'$, and $h_1 \bar{\trianglerighteq}_{k-1} \ \dots \ \bar{\trianglerighteq}_{k-1} h_l$. Moreover, from this also follows that there exists $1 \leq l' \leq l-1$ with $\{h_{l'}, h_{l'+1}\} = \{\tilde{h}^{(k-1)}, \breve{h}^{(k-1)}\}$. But then

$$h_{l'+1} \trianglerighteq_{k-1} \dots \trianglerighteq_{k-1} h' \triangleright_{k-1} h'' \trianglerighteq_{k-1} \dots \triangleright_{k-1} h_{l'},$$

which contradicts that $\tilde{h}^{(k-1)}$ and $\breve{h}^{(k-1)}$ are not comparable under $\triangleright_{k-1}$.

Case 2: $h'' \triangleright_{k-1} h'$.

This case proceeds analogously to the previous one by just reversing the roles of h' and h''.

Case 3: h' and h'' are not comparable under $\trianglerighteq_{k-1}$.

If h' and h'' are equivalent under $\trianglerighteq_k$ but not under $\trianglerighteq_{k-1}$, there have to be two sequences of networks $(h_1, \dots, h_l)$ and $(\bar{h}_1, \dots, \bar{h}_{\bar{l}})$ with $h_1 = \bar{h}_{\bar{l}} = h'$, $h_l = \bar{h}_1 = h''$, and

$$h_1 \bar{\trianglerighteq}_{k-1} \ \dots \ \bar{\trianglerighteq}_{k-1} h_l = \bar{h}_1 \bar{\trianglerighteq}_{k-1} \ \dots \ \bar{\trianglerighteq}_{k-1} \bar{h}_{\bar{l}}.$$

Moreover, there exist $1 \leq l' \leq l-1$ and $1 \leq \bar{l}' \leq \bar{l}-1$ with $\{h_{l'}, h_{l'+1}\} = \{\bar{h}_{\bar{l}'}, \bar{h}_{\bar{l}'+1}\} = \{\tilde{h}^{(k-1)}, \breve{h}^{(k-1)}\}$. In particular, this yields

$$h_{l'} \bar{\trianglerighteq}_{k-1} h_{l'+1} \trianglerighteq_{k-1} \dots \trianglerighteq_{k-1} h'' \trianglerighteq_{k-1} \dots \trianglerighteq_{k-1} \bar{h}_{\bar{l}'} \bar{\trianglerighteq}_{k-1} \bar{h}_{\bar{l}'+1}$$

which could only be satisfied if $\tilde{h}^{(k-1)}$ and $\breve{h}^{(k-1)}$ are comparable under $\trianglerighteq_{k-1}$. $\square$

Proof of Proposition 3.4

Let $(h_0, \dots, h_k)$ with $h_0, \dots, h_k \in H$ be an improving path. Moreover, suppose $h_0 = h_k$, that is, suppose $\{h_0, \dots, h_k\}$ forms a cycle. By construction of improving paths there exists $c_0 \in M$ and $D_0 \in \mathcal{D}_{h_0}^{c_0}$ with $h_1 = h_0 \pm (c_0, D_0)$.

Case 1: $D_0 \nsubseteq h_0(c_0)$, i.e., there exists $i_0 \in D_0 \setminus h_0(c_0)$.

Note that this implies $h_1 \succ^{i_0} h_0$. Moreover, since all players are self-concerned, we get

$$h_1 \succ^{i_0} h_0 \sim^{i_0} h_0 \pm (c_0, D_0 \setminus \{i_0\}) = h_1 - (c_0, \{i_0\}).$$

In other words, after joining the connection, player i_0 has no incentive to leave it unilaterally. By equability this is also true for all other $i \in h_1(c_0)$. Let $D \in \mathcal{D}_{h_1}^{c_0}$ with $D \cap h_1(c_0) \neq \varnothing$ and let $i \in D \cap h_1(c_0)$. Then:

$$h_1 \succ^i h_1 - (c_0, \{i\}) \sim^i (h_1 - (c_0, \{i\})) \pm (c_0, D \setminus \{i_0\}) = h_1 \pm (c_0, D)$$

Since the constitutions satisfy protection against eviction by assumption, no player can be forced to leave a connection against her will. Thus, all players in $h_1(c_0) \cap D$ would block deviating from h_1 to $h_1 \pm (c_0, D)$. It will be shown next that the same is also true in h_2. To this end, let $c_1 \in M$ and $D_1 \in \mathcal{D}_{h_1}^{c_1}$ with $h_2 = h_1 \pm (c_1, D_1)$. If $c_1 = c_0$, the previous discussion implies $D_1 \cap h_1(c_0) = \varnothing$ and, by similar arguments as before, it can be shown that $h_2 \succ^i h_2 \pm (c_0, D)$ for all $i \in h_2(c_0)$ and $D \in \mathcal{D}_{h_2}^{c_0}$ with $i \in D$. However, if $c_1 \neq c_0$, then $h_2(c_0) = h_1(c_0)$. Thus, by equability $h_2 \succ^i h_2 \pm (c_0, D)$ for all $i \in h_2(c_0)$ and $D \in \mathcal{D}_{h_2}^{c_0}$ with $i \in D$. Iterating these arguments implies

$$h_l \succ^i h_l - (c_0, D) \text{ for all } 1 \leq l \leq k,\ i \in h_l(c_0), \text{ and } D \in \mathcal{D}_{h_l}^{c_0} \text{ with } i \in D.$$

In particular, if $h_0 = h_k$, then $h_0 = h_k \succ^{i_0} h_k - (c_0, D_0) = h_1$ and, thus, i_0 would have blocked deviating to the network h_1.

Case 2: $D_0 \subseteq h_0(c_0)$, i.e., $h_1 = h_0 - (c_0, D_0)$.

Thus, $h_1(c_0) \subsetneq h_0(c_0)$ and, moreover, $h_0 - (c_0, D_0) \succ^i h_0$ for all $i \in D_0$ by protection against eviction. Let $i_0 \in D_0$. Since $h_0 = h_k$, there must be $1 \leq k' \leq k - 1$ and $D \in \mathcal{D}_{h_{k'}}^{c_0}$ with $h_{k'+1} = h_{k'} \pm (c_0, D)$ and $i_0 \in D$. Note that this implies $h_{k'+1} \succ^{i_0} h_{k'}$. Similar to Case 1, exploiting that all players are self-concerned yields

$$h_{k'+1} \succ^{i_0} h_{k'} \sim^{i_0} h_{k'} \pm (c_0, D \setminus \{i_0\}) = h_{k'} - (c_0, \{i_0\}).$$

Therefore, from equability follows $h_{k'+1} \succ^i h_{k'+1} - (c_0, \{i\})$ for all $i \in h_{k'+1}(c_0)$. Now, by advancing analog arguments as in Case 1 it is possible to show that this also implies

$$h_l \succ^i h_l - (c_0, D) \text{ for all } k' + 1 \leq l \leq k,\ i \in h_l(c_0), \text{ and } D \in \mathcal{D}_{h_l}^{c_0} \text{ with } i \in D.$$

In particular, this is true for $h_0 = h_k$, too. But this contradicts $h_0 - (c_0, D_0) = h_1 \succ^{i_0} h_0$. □

Proof of Proposition 3.5

Let $(h_0, h_1, \ldots, h_k)$ be an improving path in H. We will show by induction that there is always at least one player $i \in N$ with $M_{h_k}(i) \neq M_{h_0}(i)$ and $h_k \succ^i h_0$. Thus, $h_k \neq h_0$.

$k = 1$: According to the definition of an improving path and because all constitutions satisfy protection against eviction, at least one of the deviating players strictly benefits from moving to h_1. Thus, there remains nothing to show.

$k > 1$: Suppose the statement is true for $k-1$. Note that $M_{h_{k-1}}(i) \neq M_{h_0}(i)$ and $h_{k-1} \succ^i h_0$ implies $M_{h_{k-1}}(i) \bar{\succ}^i M_{h_0}(i)$. Let $c_{k-1} \in M$ be the connection and $D_{k-1} \in \mathcal{D}_{h_{k-1}}^{c_{k-1}}$ be the subset of players with $h_k = h_{k-1} \pm (c_{k-1}, D_{k-1})$. First consider the case $i \in D_{k-1}$. By assumption every player $j \in D_{k-1}$ strictly benefits from the deviation. Because preferences are lexicographic, this implies not only $h_k \succ^i h_0$ but also $M_{h_k}(i) \neq M_{h_0}(i)$. Next suppose $i \notin D_{k-1}$. Then, of course, $M_{h_k}(i) = M_{h_{k-1}}(i) \neq M_{h_0}(i)$. But it might be possible that i suffers from this deviation, i.e., $h_{k-1} \succ^i h_k$. Nevertheless, since $M_{h_k}(i) = M_{h_{k-1}}(i) \bar{\succ}^i M_{h_0}(i)$, the player still strictly prefers h_k to h_0. □

Some of the following proofs use similar technical arguments and the following lemma will serve as a convenient and useful tool. Recall that for each $h \in \mathcal{H}$,

$$\mathcal{A}_h^c(\mathcal{C}) = \left\{D \in \mathcal{D}_h^c \mid \exists S \in \mathcal{S}_h^c(D) \text{ such that } h \pm (c, D) \succ^i h \ \forall i \in (D \setminus h(c)) \cup S\right\}$$

is the set of all feasible deviations causing an instability in $c \in M$. We say that a network $h \in H$ is *exit-proof* if $D \in \mathcal{A}_h^c(\mathcal{C})$ implies $D \nsubseteq h(c)$ for all $c \in M$. That is, given an exit-proof network, no group of players $D \subseteq h(c)$ wants or is forced to leave a connection.

Lemma 3.2. Let $(N, M, \succeq, \mathcal{C})$ be a society. Moreover, let $h \in \mathcal{H}$ be an arbitrary network. Then there exists an exit-proof network $\bar{h} \in \mathcal{I}(h)$.

Proof. Let $c \in M$ such that there exists $D \in \mathcal{A}_h^c$ with $D \subseteq h(c)$. If such a connection does not exist, the network is exit-proof already and there remains nothing to show. Consider $h' := h - (c, D)$. If $h'(c)$ is not exit-proof, further subsets of players can be deleted from c until the connection is either empty or no coalition is supporting these deviations any more. This proceeding can be repeated for all connections and because N and M are finite, after finitely many steps an exit-proof network $\bar{h}$ will be reached. □

Note that by applying the previous result, Lemma 3.1 could be restated as follows: There exists no closed cycle if and only if, for each exit-proof network $\bar{h} \in \mathcal{H}$ that is not constitutionally stable, there exists an improving path leading from $\bar{h}$ to a constitutionally stable network.

Proof of Proposition 3.6

The main idea of the proof is to construct for every network in H an improving path leading from this network to a stable network. By closedness, this stable network is in H, too. Hence, there cannot be a closed cycle in H.
For constructing these paths, let us define, for each network $h \in H$, the set

$$\bar{M}_h = \{c \in M \mid \exists\, j \in h(c) : h \succ^j h - (c, \{j\})\}.$$

That is, a connection $c \in M$ is contained in $\bar{M}_h$ if and only if at least one of its members does not want to leave c. In particular, if this is the case, due to equability none of the members wants to leave the connection.
Let $h_1 \in H$ be an arbitrary network. By applying Lemma 3.2 we may assume that h_1 is exit-proof. In the following, we will establish that if h_1 is not constitutionally stable (if this would be the case, there would remain nothing to be shown), there exists an improving path from h_1 to another exit-proof network h_2 such that either $\bar{M}_{h_1} \subsetneq \bar{M}_{h_2}$, or $\bar{M}_{h_1} = \bar{M}_{h_2}$ and $h_1 \subsetneq h_2$. Then, if h_2 is not constitutionally stable, it is possible to iterate the previous step again and again. In particular, each time the step is iterated, either there are more connections which the corresponding members do not want to leave or the network strictly grows. Since both, the set of connections and the set of players, are supposed to be finite, this procedure will end after finitely many steps.

Case 1: There exists $c \in M \setminus \bar{M}_{h_1}$ with $\mathcal{A}^c_{h_1} \neq \varnothing$.

Note that because h_1 is exit-proof, $D \in \mathcal{A}^c_{h_1}$ if and only if $D \not\subseteq h_1(c)$, i.e., there is at least one player $i_1 \in D \setminus h_1(c)$ who joins the connection. Let $h_1' := h_1 \pm (c, D)$. Since all players are self-concerned, this implies

$$h_1' \succ^{i_1} h_1 \sim^{i_1} h_1 \pm (c, D \setminus \{i_1\}) = h_1' - (c, \{i_1\}).$$

In other words, after joining the connection, player i_1 has no incentive to leave it unilaterally. By equability this is also true for all $i \in h_1'(c)$ and, thus, $c \in \bar{M}_{h_1'}$. Now let $c' \in \bar{M}_{h_1}$. Note that $c \neq c'$ and $h_1(c') = h_1'(c')$. Therefore, equability implies that $c' \in \bar{M}_{h_1'}$, too. Moreover, assume there exists $D' \in \mathcal{A}_{h_1'}^{c'}$ with $D' \subseteq h_1'(c')$. That is, assume that c' is not exit-proof any more. Let $S' \in \mathcal{S}_{h_1'}^{c'}(D')$ be the corresponding supporting coalition. From regularity it follows that there is a player $j \in S'$ with $h_1 \succeq^j h_1 - (c', D')$ but $h_1' - (c', D') \succeq^j h_1'$. If $j \notin D'$, this would contradict separability because $h_1(c') = h_1'(c')$. If $j \in D'$, this would violate equability and self-concern. Therefore, the assumption cannot be true or, in other words, transforming c does not affect exit-proofness of c'. Similar considerations also apply if $c' \in M \setminus \bar{M}_{h_1}$ with $c' \neq c$. However, it might be possible that c itself is not exit-proof any more. In this case, we can delete (analogously to Lemma 3.2) all groups of players from the connection under the conditions that (i) no player joins c and (ii) all deviations comply with the constitutions, i.e., they are feasible and supported by a supporting coalition. Let h_2 be the network which is finally reached by means of this procedure. In particular, by advancing the same arguments as before it can be shown that the other connections are still exit-proof and, moreover, $\bar{M}_{h_1} = \bar{M}_{h_2} \setminus \{c\} \subsetneq \bar{M}_{h_2}$.

Case 2: $\mathcal{A}_{h_1}^{c} = \varnothing$ for all $c \in M \setminus \bar{M}_{h_1}$.

Since h_1 is not constitutionally stable, there exists $c_1 \in \bar{M}_{h_1}$ with $\mathcal{A}_{h_1}^{c_1} \neq \varnothing$. Let $D \in \mathcal{A}_{h_1}^{c_1}$ be of minimal size, i.e., $\tilde{D} \notin \mathcal{A}_{h_1}^{c_1}$ for all $\tilde{D} \subsetneq D$. Moreover, let $S \in \mathcal{S}_{h_1}^{c_1}(D)$ be the corresponding coalition which supports the deviation of D. We will show first that $D \cap h_1(c_1) = \varnothing$, that is, there are only players in D who join the connection c. Assume this is not true, i.e., there exists $i \in D \cap h_1(c_1)$. Then, $h_1 \succ^i h_1 - (c_1, \{i\}) \sim^i h_1 \pm (c_1, D)$ by self-concern and definition of $\bar{M}_{h_1}$. From this, it follows that i would not support the deviation of D and, thus, $S \cap D = \varnothing$. Since all constitutions are supposed to be decomposable and regular, we also have $\{i\} \in \mathcal{D}_{h_1}^{c_1}$ and $S \in \mathcal{S}_{h_1}^{c_1}(\{i\})$. By construction of h_1 the network is exit-proof and, therefore, there exists a player $j \in S$ with $h_1 \succeq^j h_1 - (c_1, \{i\})$. In particular, due to uniformity this is true for all members of S. But exploiting separability then yields $h_1 \pm (c_1, D \setminus \{i\}) \succeq^j h_1 \pm (c_1, D) \succ^j h_1$ for all $j \in S$ which contradicts minimality of D.

Define $h_2 \coloneqq h_1 + (c_1, D)$. As all $i \in D$ agreed to joining c_1 we must have $h_2 \succ^i h_1 \sim^i h_2 - (c_1, \{i\})$ by self-concern. Therefore, from equability it follows that no player in $h_2(c_1)$ wants to leave the connection unilaterally. Moreover, if $\bar{D} \in \mathcal{D}_{h_2}^{c_1}$ with $\bar{D} \cap h_2(c_1) \neq \varnothing$, then

$$h_2 \succ^i h_2 - (c_1, \{i\}) \sim^i h_2 \pm (c_1, \bar{D}) \tag{3.4}$$

for all $i \in \bar{D} \cap h_2(c_1)$, again by self-concern. In other words, all players who would have to leave the connection would suffer from this deviation.
In the remainder of the proof we will show that h_2 is indeed exit-proof. Let $c' \in M$ be an arbitrary connection and $D' \in \mathcal{D}_{h_2}^{c'}$ with $D' \subseteq h_2(c')$. Recall that $\mathcal{D}_{h_2}^{c'} = \mathcal{D}_{h_1}^{c'}$ by regularity and, thus, $D' \in \mathcal{D}_{h_1}^{c'}$, too.
First consider the case $c_1 \neq c'$. Since the agents' preferences are separable, $h_2 \succeq^j h_2 - (c', D')$ if and only if $h_1 \succeq^j h_1 - (c', D')$ for all $j \in h_2(c') \setminus D'$. Therefore, if $j \in h_2(c') \setminus D'$ does not support the deviation of D' in h_1, the same goes for h_2, too. However, this is also true for all $j \in D'$ due to equability and self-concern. From this it follows that a coalition supports a deviation in h_2 if and only if it does the same in h_1 (cf. Case 1). In particular, this implies that the connection c' is also exit-proof in h_2.
Next consider $c' = c_1$. Here we have to distinguish two cases, $S = \varnothing$ and $S \neq \varnothing$. First consider $S = \varnothing$, that is, when deviating from h_1 to h_2, the agents in D do not need the consent of other members for entering c. Assume there exists $D' \in \mathcal{A}_{h_2}^{c_1}$ with $D' \subseteq h_2(c_1)$. Let $S' \in \mathcal{S}_{h_2}^{c_1}(D')$ be a coalition which supports the deviation of D', i.e., there is no $j \in S'$ with $h_2 \succeq^j h_2 - (c_1, D')$. From Equation (3.4) we get $D' \cap S' = \varnothing$. Moreover, regularity implies that there exists $\varnothing \neq S'' \in \mathcal{S}_{h_1}^{c_1}(D')$ with $S'' \subseteq S'$. Note that $h_2 - (c_1, D') = (h_1 + (c_1, D)) - (c_1, D') = h_1 \pm (c_1, D \pm D')$. In particular, $D' \subseteq h_1(c_1)$ if and only if $D \cap D' \neq \varnothing$. However, this is not possible because this would contradict separability of the players' preferences. Therefore, $D \cap D' \neq \varnothing$. But this is not possible, too: by decomposability and regularity also $D \cap D' \in \mathcal{D}_{h_2}^{c_1} \subseteq \mathcal{D}_{h_1}^{c_1}$ and $S' \in \mathcal{S}_{h_2}^{c_1}(D \cap D')$. Since $\varnothing \in \mathcal{S}_{h_1}^{c_1}(D \cap D')$, decomposability and regularity again imply $D \cap D' \subseteq S'$ which contradicts Equation (3.4). Next consider $S \neq \varnothing$. We will show that $|D| = 1$. Let $i \in D$. If there would be no player $j \in S$ with $h_1 + (c, \{i\}) \succ^j h_1$, decomposability together with separability would imply $h_1 + (c, D \setminus \{i\}) \succeq^j h_1 + (c, D) = h_2 \succ^j h_1$ for all $j \in S$. In other words, S would also support a deviation of $D \setminus \{i\}$. Moreover, from uniformity it follows that $h_1 + (c, D \setminus \{i\}) \succeq^j h_1 + (c, D) = h_2 \succ^j h_1$ for all $j \in h_1(c) \cup (D \setminus \{i\})$. Thus, the players in $D \setminus \{i\}$ would agree to joining c without player i which would contradict minimality of D. Therefore, given that each $i \in D$ is supported by at least one player in S, from uniformity it follows that this is also true for all other members of $h_1(c_1)$. That is, $h_1 + (c_1, \{i\}) \succ^j h_1$ for all $j \in h_1(c_1)$. Thus, $h_1 + (c_1, \{i\}) \succ^j h_1 \succ^j h_1 - (c_1, \{j\}) \sim^j (h_1 + (c_1, \{i\})) - (c_1, \{j\})$ as $c_1 \in \bar{M}_{h_1}$. By equability this also holds for player i or, phrased differently, i has an

incentive for joining c_1 unilaterally. In fact, this implies $D = \{i\}$ by minimality of D. Moreover, due to uniformity all players in $h_1(c_1)$ strictly benefit from deviating from h_1 to h_2. Now let D', S', and S'' be given as in the case $S = \varnothing$. Then, as before we have $D' \cap D \neq \varnothing$ and, thus, $i \in D'$. By decomposability also $(h_1(c_1) \cap (D \pm D')) = h_1(c_1) \cap D' \in \mathcal{D}_{h_1}^{c_1}$ and $S'' \in \mathcal{S}_{h_1}^{c_1}(h_1(c_1) \cap D')$. Since we have $\bar{D} \in \mathcal{A}_{h_1}^{c_1}$ only if $\bar{D} \nsubseteq h_1(c_1)$, there exists $j \in S''$ with $h_1 \succeq^j h_1 - (c_1, h_1(c_1) \cap D')$. But this implies

$$h_1 - (c_1, h_1(c_1) \cap D') = h_2 - (c_1, D'^j h_2 \succ^j h_1 \succeq^j h_1 - (c_1, h_1(c_1) \cap D')$$

which obviously is a contradiction. Thus, the assumption $D' \in \mathcal{A}_{h_1}^{c_1}$ with $D' \subseteq h_2(c_1)$ must be false and c_1 is also exit-proof in h_2. □

Proof of Proposition 3.7

The proof proceeds in a similar way as the one of Proposition 3.6. As above we will construct for every exit-proof network $h_1 \in H$ an improving path leading to a stable network.

Step 1: In this step we establish that if h_1 is not constitutionally stable, there exists an improving path to another exit-proof network h_2 such that there is $D_1 \subseteq N$ with $h_2 \succ^i h_1$ and $M_{h_1}(i) \neq M_{h_2}(i)$ for all $i \in D_1$. Note that this implies $h_1 \neq h_2$.

Therefore, suppose h_1 is not constitutionally stable. Then there exists $c_1 \in M$ with $\mathcal{A}_{h_1}^{c_1} \neq \varnothing$. Let $D_1 \in \mathcal{A}_{h_1}^{c_1}$ be of minimal size, i.e., $\tilde{D} \notin \mathcal{A}_{h_1}^{c}$ for all $\tilde{D} \subsetneq D_1$. Moreover, let $S \in \mathcal{S}_{h_1}^{c}(D_1)$ be the corresponding coalition which supports the deviation of D_1. We will show first that $|D_1| = 1$. Note that $D_1 \nsubseteq h_1(c_1)$ because h_1 is exit-proof by assumption. Moreover, for all $i \in D_1$ there is at least one $j \in S$ with $h_1 + (c_1, \{i\}) \succ^j h_1$. If this would not be satisfied, analogously to Case 2 in the proof of Proposition 3.6 we would have $D_1 \setminus \{i\} \in \mathcal{A}_{h_1}^{c_1}$ since the constitutions are decomposable and the preferences are separable and lexicographic. But this would contradict minimality of D_1. Therefore, given that each $i \in D$ is supported by at least one player in S, from uniformity it follows that this also goes for all other members of $h_1(c_1)$ and, thus, $D_1 = \{i\}$ by minimality of D_1. Moreover, by applying uniformity all members in $h_1(c_1)$ are strictly better off if i enters the connection. Next we show that c_1 is also exit-proof in $\bar{h} := h_1 + (c_1, \{i\})$. Assume this is not true, that is, assume there exists $D' \in \mathcal{A}^{c_1}_{\bar{h}}$ with $D' \subseteq \bar{h}(c_1)$. Analogously to Case 2 in the proof of Proposition 3.6 we must have $i \in D'$ because the players' preferences are lexicographic and separable. Let $S' \in \mathcal{S}^{c_1}_{\bar{h}}(D')$ be a coalition which supports the deviation of D'. Moreover, let $S'' \in \mathcal{S}^{c_1}_{h_1}(D')$ with $S'' \subseteq S'$ be defined as in Case 2 in the proof of Proposition 3.6. Then, by advancing analog arguments as above we get

$$h_1 - (c_1, h_1(c_1) \cap D') = \bar{h} - (c_1, D') \succ^j \bar{h} \succ^j h_1 \succeq^j h_1 - (c_1, h_1(c_1) \cap D')$$

which obviously is a contradiction. Thus, the assumption $D' \subseteq \bar{h}(c_1)$ must be false and c_1 is also exit-proof in $\bar{h}$.
Now, suppose there exists $c' \neq c_1$ with $\bar{D} \in \mathcal{A}^{c'}_{\bar{h}}$ for some $\bar{D} \subseteq \bar{h}(c') = h_1(c')$ and let $\bar{S} \in \mathcal{S}^{c_1}_{\bar{h}}(\bar{D})$ be the corresponding supporting coalition. Note that $\bar{D} \cap \bar{S} \neq \varnothing$ due to separability. Moreover, let $i \in \bar{D} \setminus \bar{S}$. By decomposability and regularity also $\{i\} \in \mathcal{D}^{c'}_{\bar{h}} = \mathcal{D}^{c'}_{h_1}$ and $\bar{S} \in \mathcal{S}^{c'}_{\bar{h}}(\{i\})$. Since h_1 is exit-proof, there exists $j \in \bar{S}$ with $h_1 \succeq^j h_1 - (c', \{i\})$ and, thus, also $\bar{h} \succeq^j \bar{h} - (c', \{i\})$. Therefore, since the players' preferences satisfy uniformity, $\bar{h} \succeq^{\bar{j}} \bar{h} - (c', \{i\})$ for all $\bar{j} \in \bar{h}(c') \setminus \{i\}$. By exploiting separability this yields

$$\bar{h} - (c', \bar{D} \setminus \{i\}) \succeq^{\bar{j}} \left(\bar{h} - (c', \bar{D} \setminus \{i\})\right) - (c', \{i\}) = \bar{h} - (c', \bar{D}) \succ^{\bar{j}} \bar{h}$$

for all $\bar{j} \in \bar{S}$. Iterating this argument implies $\bar{D} \cap \bar{S} \in \mathcal{A}^{c'}_{\bar{h}}$, too, and $\bar{D} \setminus \bar{S} \notin \mathcal{A}^{c'}_{\bar{h}}$. Therefore, all players in $\bar{D} \cap \bar{S} \in \mathcal{A}^{c'}_{\bar{h}}$ strictly benefit from this deviation. Note that it might be the case that there exists $j \in \bar{h}(c') \cap D$ who is worse off after this change of the connection. However, because the preferences are lexicographic, this player still strictly prefers $\bar{h} - (c, \bar{D} \cap \bar{S})$ to h_1. By iterating these arguments all subsets of members where all players agree to deviating can be deleted from all connections. Let h_2 be the network which is finally reached by means of this procedure. In particular, because of separability and uniformity, h_2 is eviction-proof, too. Moreover, since no player has to leave a connection against her will and preferences are lexicographic, all players who deviated strictly prefer h_2 to h_1.

Step 2: In this step we show that if h_2 is not stable, there exists

(i) a sequence of non-empty subsets $D_1, D_2, \ldots, D_{k-1}$, and

(ii) a sequence of exit-proof networks $h_1, h_2, h_3, \ldots, h_k$ such that there is an improving path from h_{l-1} to h_l for all $2 \leq l \leq k$ and the following two conditions are satisfied:

(a) $h_l \succ^i h_{l'}$ for all $2 \leq l \leq k$, $1 \leq l' \leq l-1$, and $i \in D_{l-1}$.

(b) if $h_l \not\succ^i h_{l-1}$, then $M_{h_l}(i) = M_{h_{l-1}}(i)$.

In particular, (a) implies $h_k \neq h_{l'}$ for all $1 \leq l' < k$. Therefore, since there are only finitely many exit-proof networks, this sequence will stop after finitely many steps and, thus, the last one has to be stable.

We will show the existence of the sequence by means of induction. For $k = 2$ see Step 2. Consequently, let $k \geq 3$ and assume there exist $h_3, \ldots, h_k$ and $D_2, \ldots, D_{k-1}$ as defined above. Moreover, suppose h_k is not stable. Since this network is exit-proof by assumption, there exists $c_k \in M$ with $\mathcal{A}_{h_k}^{c_k} \neq \emptyset$ and $D \notin h_k(c_k)$ for all $D \in \mathcal{A}_{h_k}^{c_k}$. Let $D_k \in \mathcal{A}_{h_k}^{c_k}$ be of minimal size and construct h_{k+1} analogously to h_2 in Step 2. Similar to above, players deviate only if they have a strict incentive and $h_{k+1} \succ^i h_k$ for all $i \in D_k$. First, this implies $M_{h_k}(i) = M_{h_{k+1}}(i)$ for all $i \in N$ with $h_{k+1} \not\succ^i h_k$. Second, if $i \in D_k \cap D_{k-i}$, then clearly $h_{k+1} \succ^i h_{l'}$ for all $1 \leq l' \leq k$ due to transitivity. Therefore, let $i \in D_{k+1} \setminus D_k$. If $M_{h_k}(i) = M_{h_{l'}}(i)$ for all $1 \leq l' \leq k$, we have $h_{k+1} \succ^i h_{l'}$ for each of these networks because i's preferences are lexicographic. On the other hand, if $M_{h_k}(i) \neq M_{h_1}(i)$, let $l_1 := \min\{2 \leq l \leq k \mid M_{h_{l-1}}(i) \neq M_{h_l}(i)\}$. Note that (ii) implies $h_{l_1} \succ^i h_{l_1-1}$. Thus, from lexicography it follows that $h_{l_1} \succ^i h_{l'}$ for all $1 \leq l' \leq l_1 - 1$. Next consider $l_2 := \min\{l_1 + 1 \leq l \leq k \mid M_{h_{l-1}}(i) \neq M_{h_l}(i)\}$. By advancing analog arguments as before we get $h_{l_2} \succ^i h_{l'}$ for all $1 \leq l' \leq l_2 - 1$ and, thus, iterating the procedure yields $h_{k+1} \succ^i h_{l'}$ for all $1 \leq l' \leq k$. □

Proof of Proposition 3.8

This proof proceeds similarly as the proofs of the two previous propositions. Again, we construct for every network in $H = \{h \in \mathcal{H} \mid O \cap h(c) = \{o_c\} \ \forall \ c \in M\}$ an improving path leading from this network to a stable network. Because H is closed, this stable network has to be in H, too. Therefore, let $h_1 \in H$ be an arbitrary network. Because of Lemma 3.2 we may assume that h_1 is exit-proof. Moreover, let $c_1 \in M$ be an arbitrary connection with $\mathcal{A}_{h_1}^{c_1} \neq \emptyset$. The construction of the path proceeds in three steps:

Step 1: We establish that there exists $B_1 \in \mathcal{A}_{h_1}^{c_1}$ with $\mathcal{A}_{h_1+(c_1,B_1)}^{c_1} = \emptyset$.

The main idea of this step is to exploit separability of the owner's preferences. Define $B_1 \subseteq N$ via

$$B_1 := \left\{ i \in E \setminus h_1(c_1) \;\middle|\; h_1 + (c_1, \{i\}) \succ^i h_1 \text{ and } h_1 + (c_1, \{i\}) \succ^{o_{c_1}} h_1 \right\}.$$

That is, B_1 contains exactly those players who want to join c_1 and would be accepted by o_{c_1}. Let $i, j \in B_1$. Then, $h_1 + (c_1, \{i, j\}) \succ^{o_{c_1}} h_1 + (c_1, \{i\}) \succ^{o_{c_1}} h_1$ by separability of o_{c_1}'s preferences. Iterating this argument implies $h_1 + (c_1, B_1) \succ^{o_{c_1}} h_1$. Moreover, since the workers' preferences are lexicographic, also $h_1 + (c_1, B_1) \succ^i h_1$ for all $i \in B_1$. Indeed, this yields $B_1 \in \mathcal{A}^{c_1}_{h_1}$. Now suppose there exists a deviation $D \in \mathcal{A}^{c_1}_{h_1+(c_1,B_1)}$. If $D \subseteq h_1(c_1) \cup B_1$, the definition of B_1 and exit-proofness of h_1 imply $h_1 + (c_1, B_1) \succeq^{o_{c_1}} (h_1 + (c_1, B_1)) - (c_1, i)$ for all $i \in h_1(c_1) \cup B_1$. Advancing the same arguments as before yields $h_1 + (c_1, B_1) \succeq^{o_{c_1}} (h_1 + (c_1, B_1)) - (c_1, D)$, which implies that o_{c_1} would not support the deviation. Moreover, the workers in $h_1(c_1) \cup B_1$ obviously do not want to leave the firm and, thus, $D \subseteq h_1(c_1) \cup B_1$ cannot be true. However, if $D \not\subseteq h_1(c_1) \cup B_1$ and there exists $i \in D \setminus h_1(c_1)$ with $(h_1 + (c_1, B_1)) + (c_1, \{i\}) \succ^{o_{c_1}} h_1 + (c_1, B_1)$, then by construction of B_1 and because i's preferences are lexicographic, this worker would not agree to joining c_1. Therefore, $\mathcal{A}^{c_1}_{h_1+(c_1,B_1)}$ must be empty.

Step 2: We construct an improving path leading from $h_1' := h_1 + (c_1, B_1)$ to another exit-proof network h_2 with $h_2 \succ^i h_1$ for all $i \in B_1$ and $h_2 \succeq^i h_1$ for all $i \in E \setminus B_1$. Let $c' \in M$ such that there exists $B' \subseteq h_1'(c')$ with $B' \in \mathcal{A}^{c'}_{h_1'}$ and choose B' maximal with respect to "$\subseteq$", i.e., there exists no $\bar{B} \subseteq h_1'(c')$ with $\bar{B} \in \mathcal{A}^{c'}_{h_1'}$ and $B' \subsetneq \bar{B}$. Note that $c' \neq c_1$ as $\mathcal{A}^{c_1}_{h_1'} = \emptyset$. By assumption $o_{c'}$'s preferences are separable and, thus, $h_1' \succeq^{o_{c'}} h_1' - (c', B')$ by exit-proofness of h_1. Therefore, $h_1' - (c', B') \succ^j h_1'$ for all $j \in B'$. Now suppose there exists $i \in B' \setminus B_1$. Note that $i \in h_1(c)$ if and only if $i \in h_1'(c)$ for all $c \in M$. If i has a strict incentive for leaving c' in h_1', she would also have a strict incentive for leaving the connection in h_1 because her preferences are lexicographic. But this contradicts exit-proofness of h_1 and, thus, $B' \subseteq B_1$. Moreover, by construction of B' and separability of $o_{c'}$'s preferences, there exists no further set of workers $B'' \subseteq h_1'(c') \setminus B'$ with $B'' \in \mathcal{A}^{c'}_{h_1'-(c',B')}$. By iterating the previous procedure, it is possible to reach an exit-proof network h_2 by deleting all workers from all connections they want to leave without impairing the other workers in $E \setminus B_1$. In particular, for all $i \in E \setminus B_1$ nothing changes and, therefore, they are indifferent between h_2 and h_1. On the other hand, all $i \in B_1$ strictly benefit from the deviations and, thus, they strictly prefer h_2 to h_1.

Step 3: Iterating the procedure.
Given h_2, if $A^c_{h_2} = \emptyset$ for all $c \in M$, there remains nothing to show. Therefore, assume there exists $c_2 \in M$ with $A^{c_2}_{h_2} \neq \emptyset$. By repeating Steps 2 and 3 it is pos-

sible to find $B_2 \subseteq E \setminus h_2(c_2)$ with $\mathcal{A}^{c_2}_{h_2+(c_2,B_2)} = \varnothing$ and to construct an improving path leading from $h_2 + (c_2, B_2)$ to an exit-proof network h_3. Analogously, h_2 will be Pareto dominated by h_3 from the workers' perspective. As H is finite, there exist only finitely many exit-proof networks. Hence, this procedure will end after finitely many steps. □

References

Acemoglu, D., Egorov, G. and Sonin, K. (2012). Dynamics and stability of constitutions, coalitions, and clubs. *American Economic Review*, **102** (4), 1446–1476.

Allouch, N. and Wooders, M. (2008). Price taking equilibrium in economies with multiple memberships in clubs and unbounded club sizes. *Journal of Economic Theory*, **140**, 246–278.

Bala, V. and Goyal, S. (2000). A noncooperative model of network formation. *Econometrica*, **68** (5), 1181–1229.

Banerjee, S., Konishi, H. and Sönmez, T. (2001). Core in a simple coalition formation game. *Social Choice and Welfare*, **18**, 135–153.

Bogomolnaia, A. and Jackson, M. O. (2002). The stability of hedonic coalition structures. *Games and Economic Behavior*, **38**, 201–230.

Cesco, J. C. (2012). Hedonic games related to many-to-one matching problems. *Social Choice and Welfare*, **39** (4), 737–749.

Chung, K.-S. (2000). On the existence of stable roommate matchings. *Games and Economic Behavior*, **33**, 206–230.

Diamantoudi, E., Miyagawa, E. and Xue, L. (2004). Random paths to stability in the roommate problem. *Games and Economic Behavior*, **48**, 18–28.

Dutta, B. and Massó, J. (1997). Stability of matchings when individuals have preferences over colleagues. *Journal of Economic Theory*, **75**, 464–475.

— and Mutuswami, S. (1997). Stable networks. *Journal of Economic Theory*, **76**, 322–344.

EICHENIQUE, F. and YENMEZ, M. B. (2007). A solution to matching with preferences over colleagues. *Games and Economic Behavior*, **59**, 46–71.

ELLICKSON, B., GRODAL, B., SCOTCHMER, S. and ZAME, W. (1999). Clubs and the market. *Econometrica*, **67**, 1185–1217.

—, —, — and — (2001). Clubs and the market: Large finite economies. *Journal of Economic Theory*, **101** (01), 40–77.

FARRELL, J. and SCOTCHMER, S. (1988). Partnerships. *The Quarterly Journal of Economics*, **103**, 279–297.

FERSHTMAN, C. and PERSITZ, D. (2012). Clubs and social networks. *Mimeo*.

GRIMMETT, G. R. and STIRZAKER, D. R. (2001). *Probability and Random Processes*. Oxford University Press.

HART, S. and MAS-COLELL, A. (1989). Potential, value, and consistency. *Econometrica*, **57** (03), 589–614.

JACKSON, M. O. (2008). *Social and Economic Networks*. Princeton University Press.

— and WATTS, A. (2001). The existence of pairwise stable networks. *Seoul Journal of Economics*, **14** (3), 299–322.

— and — (2002). The evolution of social and economic networks. *Journal of Economic Theory*, **106**, 265–295.

— and WOLINSKY, A. (1996). A strategic model of social and economic networks. *Journal of Economic Theory*, **71** (0108), 44–74.

JEHIEL, P. and SCOTCHMER, S. (2001). Constitutional rules of exclusion in jurisdiction formation. *The Review of Economic Studies*, **68**, 393–413.

KANDORI, M., MAILATH, G. J. and ROB, R. (1993). Learning, mutation, and long run equilibria in games. *Econometrica*, **61** (1), 29–56.

— and ROB, R. (1995). Evolution of equilibria in the long run: A general theory and applications. *Journal of Economic Theory*, **65** (2), 383–414.

Klaus, B., Klijn, F. and Walzl, M. (2010). Stochastic stability for roommate markets. *Journal of Economic Theory*, **145**, 2218–2240.

Kojima, F. and Ünver, M. U. (2008). Random paths to pairwise stability in many-to-many matching problems: a study on market equilibration. *International Journal of Game Theory*, **36**, 473–488.

Kominers, S. D. (2010). Matching with preferences over colleagues solves classical matching. *Games and Economic Behavior*, **68**, 773–778.

Monderer, D. and Shapley, L. S. (1996). Potential games. *Games and Economic Behavior*, **14** (44), 124–143.

Myerson, R. B. (1980). Conference structures and fair allocation rules. *International Journal of Game Theory*, **9** (3), 169–182.

Page, F. H. and Wooders, M. H. (2009). Strategic basins of attraction, the path dominance core, and network formation games. *Games and Economic Behavior*, **66**, 462–487.

— and — (2010). Club networks with multiple memberships and noncooperative stability. *Games and Economic Behavior*, **70**, 12–20.

—, — and Kamat, S. (2005). Networks and farsighted stability. *Journal of Economic Theory*, **120** (2), 257–269.

Pycia, M. (2012). Stability and preference alignment in matching and coalition formation. *Econometrica*, **80** (1), 323–362.

Qin, C.-Z. (1996). Endogenous formation of cooperation structures. *Journal of Economic Theory*, **69** (0047), 218–226.

Roth, A. E. (1984). Stability and polarization of interests in job matching. *Econometrica*, **52** (1), 47–58.

— (1985). The college admissions problem is not equivalent to the marriage problem. *Journal of Economic Theory*, **36**, 277–288.

— and Sotomayor, M. (1990). *Two-sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.

— and Vande Vate, J. H. (1990). Random paths to stability in two-sided matching. *Econometrica*, **58**, 1475–1480.

Slikker, M. (2001). Coalition formation and potential games. *Games and Economic Behavior*, **37**, 436–448.

Sotomayor, M. (1999). Three remarks on the many-to-many stable matching problem. *Mathematical Social Science*, **38**, 55–70.

— (2004). Implementation in the many-to-many matching market. *Games and Economic Behavior*, **46**, 199–212.

Chapter 4

Robust Equilibria in Location Games

4.1 Introduction

In his classic example, Harold Hotelling illustrates competition in a heterogeneous market by two firms that consider where to place their shop on a main street (Hotelling, 1929). Ever since, this model of spatial competition has inspired a tremendous amount of research in various disciplines. Starting with Downs (1957), it is used to analyze the positioning of political candidates competing for voters (e.g., Mueller, 2003; Roemer, 2001) and to analyze the positioning of products in order to attract consumers (e.g., Carpenter, 1989; Salop, 1979). In the year 2013 alone, Hotelling has been cited more than 450 and Downs even more than 1100 times.[1] Moreover, the model implication of *minimal differentiation* is known far beyond scholarly circles. In this paper we want to challenge a fundamental aspect of the Hotelling-Downs approach.

Throughout the literature (of spatial competition), it has been virtually always assumed that consumers or voters who prefer the same position fully agree upon the ranking of the other alternatives, i.e., they have identical preferences or utility functions. This very strong homogeneity requirement can be considered as driven by the assumption that all consumers/voters use the same distance measure since in the standard Hotelling-Downs set-up (dis)utility is represented by the distance between positions. In particular, if two people prefer the same option, in any spatial represen-

[1] Google Scholar, February 10, 2014.

tation with homogeneous distances they necessarily rank all the other alternatives in the same order. This is hard to justify when we think of voters of the same political party who disagree about the second-best party, or of consumers with the same favorite brand but disagreement about the ordering of two other brands. Even in the case of geographic location choices the requirement appears to be challengeable if distances represent travel time, for instance.[2] As a matter of fact, these simple cases already exceed the scope of almost any model of locational competition.

Consider, for example, a poll on a group of voters about their favorite tax rate. The answers can be displayed as locations on a line. Location games that capture this application consider classically two political candidates who strategically choose a tax rate which they propose to the voters. Thereby it is standardly assumed that (a) each voter casts his vote for the candidate that is closest to him and (b) all voters asses the distances between the candidates homogeneously. In combination these two assumptions are not at all innocuous. As indicated above, they hide the homogeneity requirement that all voters who consider a tax rate of 10 percent, for instance, as their favorite alternative, are supposed to rank any two tax rates, like 2 percent and 20 percent, for example, in exactly the same order. Since this requirement is unnaturally strong, the classical result that two vote maximizing candidates choose the median location (Hotelling, 1929) stands apparently on highly questionable grounds. A way to avoid this issue would be to ask the participants in the poll not only about their favorite tax rate, but about a full ranking of the alternative tax rates. Apart from practical problems, the downside of such an approach is the informational requirement that political candidates know the full assessment of every voter. That is, we have replaced a questionable requirement by another one. A solution to this issue relates back to the seminal contribution of Black (1948). He examined single-peaked preferences on a line, which has the same effect as voters who are allowed to asses the "distances" between different tax rates individually. Black's result that under single-peaked preferences the median voter wins in majority voting against any other alternative has the following implication for the situation of spatial competition outlined above: In *any* location game that is consistent with the poll, both candidates choose the median tax rate in equilibrium. In that sense the classical result is *robust*.

[2]Indeed, it is possible that two individuals differ in their speed of walking uphill such that they would not choose the same path although both easily agree that there is one short and steep path and one longer and flatter path.

The example on tax rates illustrates that in two-player location games on a line the questionable requirement of homogeneous distance perceptions is not driving the final outcome. However, for all other cases – in particular, for more than two players and for multidimensional spaces – robustness of the results is an open problem. If one can show that the model assumption is not driving the results, then the model is put on a solid foundation. This issue, although fundamental, seems to have been overlooked in the – rich and exciting – history of location games.
In this paper we want to scrutinize for given outcomes of spatial competition whether they rely on homogeneous distance perceptions or not. To this end, we formalize individual distance perceptions as individual edge lengths of a graph.[3] A formal description of consumers/voters of this type leads to a non-cooperative game between p players, which are interpreted as firms or political candidates. In this game, players simultaneously choose a location in order to maximize the number of agents (i.e., consumers/voters) they can attract. An equilibrium is then called *robust* if it is an equilibrium for all possible distance perceptions that are based on the same underlying structure (a line, for example). In other words, our modeling approach boils down to defining a stronger notion of equilibrium which we call *robust equilibrium*. It is defined directly on the situation of spatial competition, i.e., the underlying space and the distribution of agents (such as the poll on tax rates). Formally, several of location games correspond to the same situation of spatial competition, one for each setting of individual distance perceptions; and a robust equilibrium is a Nash equilibrium in any of these games. In particular, it is also a Nash equilibrium in the standard case of homogeneous distances.

A key result for our analysis is the characterization of robust equilibria by four conditions which are jointly necessary and sufficient. It is based on partitioning the underlying space into "hinterlands" and "competitive zones". Applying this result allows us first of all to judge which of the standard results are robust. In fact, we find that several outcomes do not depend on the assumption of homogeneous distances, but others do.
In the second part of the paper we examine general properties of robust equilibria.

[3]This can be shown to be equivalent to the assumption of single-peaked preferences on certain domains. For example, if the underlying structure is a line graph, then this assumption is equivalent to the standard notion of single-peakedness. An alternative model variation would keep the assumption of homogenous distances but add a set of nodes (which we call "dummy nodes") to make the graph more flexible. However this model variation can be shown to undermine the model's explanatory power.

Among them is the central issue of minimal differentiation (e.g., d'Aspremont *et al.*, 1979; de Palma *et al.*, 1985, 1990; Eaton and Lipsey, 1975; Economides, 1986; Król, 2012; Meagher and Zauner, 2004). It turns out that robust equilibria satisfy a local variant of minimal differentiation, i.e., they induce reduced games in which the corresponding players are minimally differentiated. This result provides strong support for the "principle of minimal clustering" which has been proposed in the seminal contribution of Eaton and Lipsey (1975). Indeed, for any number of players, any underlying structure, and any distribution of agents, robust equilibria are characterized by clusters of players. That is, the players are jointly located on what we show to be the appropriately defined medians of local areas. Based on this result, we discuss the welfare implications for consumers and observe that almost all robust equilibria are not Pareto efficient. Consumers would unambiguously improve if some firm would be relocated appropriately. We finally, elaborate on the conditions for the existence of robust equilibria. We analyze how the spatial structure and the distribution of consumers/voters guarantee, admit, or preclude the existence of robust equilibria. Interestingly, two very common assumptions in the literature – (a) uniform distribution of consumers/voters and (b) one-dimensional space such as cycle or line structures – are mutually exclusive in the sense that for higher numbers of players robust equilibria require that one of them is not satisfied.

Related Literature

There is an immense body of literature on spatial competition. While the original Hotelling-Downs framework is restricted to a one-dimensional space, a uniform distribution of agents, and only two players, many authors have attempted to relax these restrictions. To do so, one branch of the literature has followed a continuous modeling approach within the Euclidean space $\mathbb{R}^k$ (e.g., d'Aspremont *et al.*, 1979; Economides, 1986), while a second branch replaces the Euclidean space by a graph (e.g., Labbé and Hakimi, 1991). Because the history of both branches is rich and long, providing a summary which covers all of it would exceed the scope of our paper. We restrict ourselves here to list several surveys on the topic and to discuss the most closely related works.

A broad overview and taxonomy of literature on spatial competition can be found in Eiselt *et al.* (1993). Based on five components (the underlying space, the number of players, the pricing policy, the rules of the game, and the behavior of the agents) the authors provide a bibliography for competitive location models. While this sum-

mary is not limited to certain subbranches, more specific surveys have been written on spatial models of consumer product spaces (Lancaster, 1990), on spatial competition in continuous space (Gabszewicz and Thisse, 1992), on spatial models of political competition (Mueller, 2003; Osborne, 1995), on competition in discrete location models (Plastria, 2001), on sequential competition (Eiselt and Laporte, 1997; Kress and Pesch, 2012), and on one-stage competition in location models (Eiselt and Marianov, 2011; ReVelle and Eiselt, 2005).
Although there are many variations and relaxations of spatial competition, virtually all of the models rely on the assumption of homogeneous distance perceptions. For instance, asymmetric transportation costs (e.g., Nilssen, 1997) do not alter the assumption. In order to examine to which extent this standard simplification is driving the results we will focus on the first stage of Hotelling's game, i.e., we will investigate the location choices of the players but we will not include additional variables such as prices. Similar approaches have been used, for example, by Eaton and Lipsey (1975), Denzau *et al.* (1985), and Braid (2005) who also concentrate on spatial competition by assuming fixed (and equal) prices. Nevertheless, extending our approach to a two-stage game would be a potential next step for further research. Integrating heterogeneous consumer behavior into a model of spatial competition has been attempted by a few studies only. Among them are de Palma *et al.* (1985, 1990) and Rhee (1996) who find that ambiguity about consumers' (or voters') behavior may lead to minimal differentiation. More specifically, they show that if the consumers' preferences do not only depend on prices and distances but also on inherent product characteristics and, furthermore, the firms have incomplete information about consumers' tastes, then Hotelling's main result can be restored under certain conditions. This conclusion is not confirmed in closely related models where the authors assume that the exact position of demand is unknown (e.g. Król, 2012; Meagher and Zauner, 2004, 2005). Thus, the validity of minimal differentiation under heterogeneous agents is still an open problem and the same holds true for the main implications, like that spatial competition generically does not lead to socially efficient outcomes, for example. However, the previously cited publications differ from our work in at least two important aspects. First, in these works, players are assumed to have a probability distribution for the behavior of agents. In our work, uncertainty is not explicitly modeled but only enters implicitly as robust equilibria do not depend on specification details about the agents' behavior. Second, the way we model and interpret heterogeneity differs from the approaches of the other authors. In our setting, the agents apply individual distances to compare specific

product variations but the preferences do not depend on inherent product characteristics. To model this in a convenient way we use a graph-based approach. We believe that our definitions are more intuitive in discrete spaces than in the plane and that this approach helps to highlight the difference between homogeneous and heterogeneous agents. However, the main questions of our work are not restricted to graphs and thus our contribution should also be interesting in a more general context. To the best of our knowledge, this is the first paper that assesses robustness of equilibria in location games with respect to different distance perceptions.
From a technical point of view, the model from Eiselt and Laporte (1991, 1993) is heavily related to ours. In these publications the authors show for homogeneous agents that the two-player and three-player cases on trees always result in some kind of minimal differentiation. We will check whether this is also true in our more general context of more than two players and arbitrary graphs. More recently, Shiode and Drezner (2003) studied the two-player case on trees under sequential location choices and stochastic demand. Further recent contributions, to name but a few, deal with terrorism (e.g., Berman and Gavious, 2007) or stem from computer science (e.g., Godinho and Dias, 2010; Jiang *et al.*, 2011; Mavronicolas *et al.*, 2008). Still, the issue of heterogeneous distances has not been addressed in any of these publications.

Related Literature

There is an immense body of literature on spatial competition. While the original Hotelling-Downs framework is restricted to a one-dimensional space, a uniform distribution of agents, and only two players, many authors have attempted to relax these restrictions. To do so, one branch of the literature has followed a continuous modeling approach within the Euclidean space $\mathbb{R}^k$ (e.g., d'Aspremont *et al.*, 1979; Economides, 1986), while a second branch replaces the Euclidean space by a graph (e.g., Labbé and Hakimi, 1991). Because the history of both branches is rich and long, providing a summary which covers all of it would exceed the scope of our paper. We restrict ourselves here to list several surveys on the topic and to discuss the most closely related works.
A broad overview and taxonomy of literature on spatial competition can be found in Eiselt *et al.* (1993). Based on five components (the underlying space, the number of players, the pricing policy, the rules of the game, and the behavior of the agents) the authors provide a bibliography for competitive location models. While this sum-

mary is not limited to certain subbranches, more specific surveys have been written on spatial models of consumer product spaces (Lancaster, 1990), on spatial competition in continuous space (Gabszewicz and Thisse, 1992), on spatial models of political competition (Mueller, 2003; Osborne, 1995), on competition in discrete location models (Plastria, 2001), on sequential competition (Eiselt and Laporte, 1997; Kress and Pesch, 2012), and on one-stage competition in location models (Eiselt and Marianov, 2011; ReVelle and Eiselt, 2005).
Although there are many variations and relaxations of spatial competition, virtually all of the models rely on the assumption of homogeneous distance perceptions. For instance, asymmetric transportation costs (e.g., Nilssen, 1997) do not alter the assumption. In order to examine to which extent this standard simplification is driving the results we will focus on the first stage of Hotelling's game, i.e., we will investigate the location choices of the players but we will not include additional variables such as prices. Similar approaches have been used, for example, by Eaton and Lipsey (1975), Denzau *et al.* (1985), and Braid (2005) who also concentrate on spatial competition by assuming fixed (and equal) prices. Nevertheless, extending our approach to a two-stage game would be a potential next step for further research. Integrating heterogeneous consumer behavior into a model of spatial competition has been attempted by a few studies only. Among them are de Palma *et al.* (1985, 1990) and Rhee (1996) who find that ambiguity about consumers' (or voters') behavior may lead to minimal differentiation. More specifically, they show that if the consumers' preferences do not only depend on prices and distances but also on inherent product characteristics and, furthermore, the firms have incomplete information about consumers' tastes, then Hotelling's main result can be restored under certain conditions. However, this conclusion is not confirmed in closely related models where the authors assume that the exact position of demand is unknown (e.g. Król, 2012; Meagher and Zauner, 2004, 2005). Thus, the validity of minimal differentiation under heterogeneous agents is still an open problem and the same holds true for the main implications, like that spatial competition generically does not lead to socially efficient outcomes, for example. The previously cited publications differ from our work in at least two important aspects. First, in these works, players are assumed to have a probability distribution for the behavior of agents. In our work, uncertainty is not explicitly modeled but only enters implicitly as robust equilibria do not depend on specification details about the agents' behavior. Second, the way we model and interpret heterogeneity differs from the approaches of the other authors. In our setting, the agents apply individual distances to compare specific product

variations but the preferences do not depend on inherent product characteristics. To model this in a convenient way we use a graph-based approach. We believe that our definitions are more intuitive in discrete spaces than in the plane and that this approach helps highlight the difference between homogeneous and heterogeneous agents. However, the main questions of our work are not restricted to graphs and thus our contribution should also be interesting in a more general context. To the best of our knowledge, this is the first paper that assesses robustness of location games with respect to different distance perceptions.

From a technical point of view, the model from Eiselt and Laporte (1991, 1993) is heavily related to ours. In these publications the authors show for homogeneous agents that the two-player and three-player cases on trees always result in some kind of minimal differentiation. We will check whether this is also true in our more general context of more than two players and arbitrary graphs. More recently, Shiode and Drezner (2003) studied the two-player case on trees under sequential location choices and stochastic demand. Further recent contributions also deal with terrorism (e.g., Berman and Gavious, 2007) or stem from computer science (e.g., Godinho and Dias, 2010; Jiang *et al.*, 2011; Mavronicolas *et al.*, 2008). Still, the issue of heterogeneous distances is not addressed in these publications.

4.2 The Model

Our modeling approach proceeds in two steps. First we consider, as usual, a non-cooperative game between players (the firms/candidates) who are able to occupy a position or object, respectively. The agents (consumers/voters) are still attracted by the player(s) located closest to them but now their distance perceptions may be assessed on an individual basis. More specifically, the agents agree on the underlying space which is modeled by means of a graph (Subsection 4.2.1), but in our setting they may individually measure the similarity between the objects (Subsection 4.2.2). Then, in the second step, we study whether equilibria of the game are robust with respect to perturbations of the distance perceptions. To this end, roughly speaking, we fully abandon the distances. This means formally that an outcome is called robust if it is an equilibrium for all possible edge lengths of the same underlying graph (Subsection 4.2.3). If this is satisfied, the outcome is completely independent of individual distance perceptions and then the standard case of homogeneous distances is a well-justified simplification.

4.2.1 Definitions of Graphs

An undirected graph (X, E) consists of a set of *vertices* or *nodes* X and a set of *edges* E where each edge is a subset of the vertices of size two. Let X be a finite set of size $\xi \geq 2$. For brevity we write xy or yx for an edge $\{x, y\} \in E$. Given a graph (X, E), we denote by $N_x := \{y \in X \mid xy \in E\}$ the set of *neighbors* of a node x. The number of edges/neighbors is its degree $\deg_x := |N_x|$. Furthermore, $Y \subseteq X \setminus \{x\}$ is neighboring to $x \in X$ if there exists some $y \in Y$ with $xy \in E$.
A *path* from $x \in X$ to $x' \in X$ in (X, E) is a sequence of distinct nodes $(x_1, \ldots, x_T)$ such that $x_1 = x$, $x_T = x'$, and $x_t x_{t+1} \in E$ for all $t \in \{1, \ldots, T-1\}$. A set of nodes $Y \subseteq X$ is said to be connected if for any pair $y, y' \in Y$ there exists a path between the two nodes. A set of connected nodes is called a *component* if there is no path to nodes outside of this set, i.e., $C \subseteq X$ is a *component* of (X, E) if it is connected and for all x, x' such that $x \in C$ and $x' \in X \setminus C$ there does not exist any path. A graph that consists of only one component is called connected because then there is a path between any two nodes. Throughout the paper, we will restrict attention to connected graphs. An important class of such graphs is the class of trees. Trees are connected with $\xi - 1$ edges or, equivalently, in a tree each pair of vertices is connected by a unique path.
A node-weighted graph is a triple (X, E, w), where $w := (w_x)_{x \in X} \in \mathbb{R}_+^{\xi}$ is a vector of weights. We write w_x for the weight of node $x \in X$ and $w(Y) = \sum_{y \in Y} w_y$ for the weight of a set of nodes $Y \subseteq X$. The weight w will be determined later on by the distribution of agents.

Now let (X, E, w) be given. An important operation in graphs is to delete a set of nodes $Y \subseteq X$ and all involved edges: $(X, E) - Y := \left(X \setminus Y, E_{|X \setminus Y}\right)$ with $E_{|X \setminus Y} = \{xy \in E \mid x, y \in X \setminus Y\}$. This is illustrated in Figure 4.1.
The operation $(X, E) - Y$ leads to a graph with potentially several components and we denote them by $C_1^Y, C_2^Y, \ldots, C_{l_Y}^Y$ such that $w(C_1^Y) \geq w(C_2^Y) \geq \ldots \geq w\left(C_{l_Y}^Y\right)$. If $l_Y > 1$ and $|Y| = 1$, say $Y = \{x\}$, the node is called a *cut vertex* (cf., e.g., Diestel, 2005) and we write C_k^x instead of $C_k^{\{x\}}$. In this case, for the number of components it holds that it is not greater than the degree of x. A connected set of nodes $B \subseteq X$ is called a *block* if there is no cut vertex in $(X, E) - X \setminus B = (B, E_{|B})$ and B is maximal with respect to inclusion, i.e., $B \subsetneq B' \subseteq X$ implies that there exists a cut vertex in $(B', E_{|B'})$. That is, a set of nodes is a block if the induced subgraph cannot be decomposed into multiple components by deleting single nodes and it is not possible to find a larger subgraph with this feature. Note that $x \in X$ is contained in several

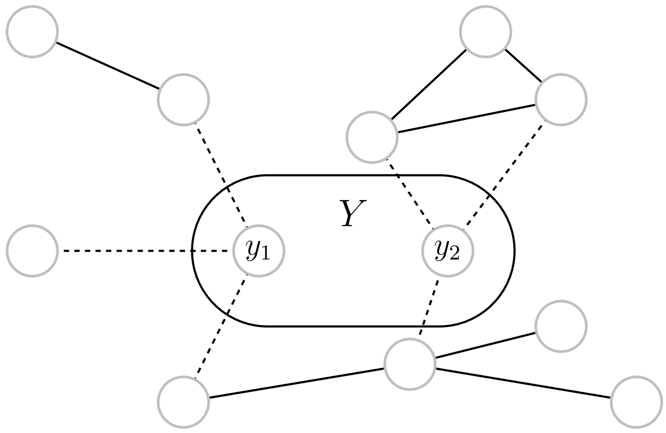


Figure 4.1: Deletion of nodes.

blocks if and only if it is a cut vertex. The set of blocks of a given graph is denoted by $\mathcal{B}$ and $b := |\mathcal{B}|$ is the number of blocks.

4.2.2 Perceived Distances and Players' Payoffs

In the following, the elements of X are called *objects* and are interpreted, according to the three applications, as geographical locations, political platforms or product specifications. Let $N = \{i_1, \ldots, i_n\}$ be a finite set of *agents* who have a favorite object $\hat{x}^i \in X$. As usual, the graph (X, E) is used to represent the relations between the objects as they are perceived by the agents. [4] In order to be as general as possible we impose no further requirements on the structure of the graph, but typical examples from literature are lines, cycles or lattices, to name but a few. In contrast to previous works, we assume that perceptions are subjective to some extent. Formally, for each $i \in N$ there are edge lengths $(\delta_e^i)_{e\in E} > 0$ that represent his individual estimation of distances between the nodes, such that, for example, δ_e^i need not coincide with δ_e^j.[5] Given $\delta := (\delta_e^i)_{e\in E}^{i\in N}$, agent i's *perceived distance* $d^i(x)$ to an object $x \in X$ is the length of the shortest path(s) from the favorite object $\hat{x}^i$ to x, where the length of a path

[4] Note that we do not allow for "dummy nodes," that is, we do not consider the possibility of adding further nodes to the graph which are not objects. This is due to the fact that dummy nodes can be shown to undermine the explanatory power of the model.

[5] The interpretation for geographic locations is as follows: The agents agree on the underlying graph (a road map, for example) but they are heterogeneous in terms of assessing or evaluating the edge lengths (the travel time, for example). If the graph does not represent geographic distances, but policy spaces or the perception of brands, it seems to be an even more unrealistic assumption that all agents use the same distance measure, as motivated in the introduction.

is the sum of its edge lengths:

$$d^i(x) := \min\left\{ \sum_{t=1}^{T-1} \delta^i_{x_t x_{t+1}} \;\middle|\; (x_1, \ldots, x_T) \text{ is a path from } \hat{x}^i \text{ to } x \right\}.$$

We set $d^i(\hat{x}^i) = 0$ for all $i \in N$. Note that two agents with the same favorite object, i.e., $\hat{x}^i = \hat{x}^j$, might have different perceptions about the distances to the other objects. As usual, we will assume a "distance-based behavior"of the agents, i.e., agent $i \in N$ weakly prefers an object $x \in X$ over $y \in X$ if and only if $d^i(x) \leq d^i(y)$. In other words: his utility is decreasing in distances. Thus, the preferences of agent $i \in N$ are completely determined by his favorite object $\hat{x}^i$ and his individual edge lengths $(\delta^i_e)_{e \in E}$.[6] With the assumption that $\delta^i_e = \delta^j_e$ for all $i, j \in N$ and any $e \in E$, we obtain the standard model, where distance perceptions are homogeneous.

In addition to the objects and agents, we consider a set of *players* $P := \{c_1, \ldots, c_p\}$ of finite size $p \geq 2$. To ease the distinction between agents and players we will use the male form for agents, while players are assumed to be female. Each $c \in P$ is supposed to occupy an object $x \in X$. Formally, the *strategy set* for each player $c \in P$ is $S^c = X$, such that a strategy is an object $s^c \in X$. Let $S = S^{c_1} \times \ldots \times S^{c_p}$. Given a strategy profile $s \in S$, let $p_x \in \mathbb{N}$ be the number of players whose strategy is $x \in X$. Furthermore, let $\Phi^i(s)$ be the set of players who are perceived as closest by agent $i \in N$, i.e., $\Phi^i(s) = \{c \in P \mid d^i(s^c) \leq d^i(s^{\bar{c}}) \;\; \forall\, \bar{c} \in P\}$. Note that we loosely speak about the perceived distance to a player $c \in P$ instead of the distance to the player's chosen object $s^c \in X$. We assume that each agent is allocated to the player which is perceived as closest. If multiple players are perceived as closest by some agent, then he is assumed to be uniformly distributed among these players. Thus, given a strategy profile $s \in S$, player c's payoff $\Phi^c(s)$ is the mass of agents who perceive object s^c as closest to their favorite object, i.e., the payoff of $c \in P$ is given by $\pi^c(s) = \sum_{i: c \in \Phi^i(s)} \frac{1}{|\Phi^i(s)|}$. A profile of payoffs is denoted by $\pi_\delta := \left(\pi^c_\delta\right)^{c \in P} := (\pi^c)^{c \in P}$, where the subscript δ indicates that the payoffs depend on the individual edge lengths $\delta = (\delta^i_e)^{i \in N}_{e \in E}$.

[6] There is a justification for this type of preference which neither deals with differing edge lengths nor with distance-based behavior. Agents can be assumed to have single-peaked preferences on the graph as they were defined for lines (Black, 1948) or trees (Demange, 1982). Such preferences find broad acceptance and play a crucial role in the literature on social choice (see, e.g., Moulin, 1980). The alternative formulation with single-peaked preferences is, in fact, equivalent to the (quite different) formulation here. The proof for this claim can be requested from the authors.

4.2.3 Equilibrium Notions

Fix a graph (X, E) and a set of agents N such that for each agent $i \in N$ we have a favorite object $\hat{x}^i \in X$ and individually measured edge lengths $(\delta_e^i)_{e \in E}$. Then a normal form game is given by $\Gamma^\delta = (P, S, \pi_\delta)$. The game is indexed by δ to emphasize that the payoffs, and therefore the game depends on the individual edge lengths. The main goal of our work is to examine to which extent this restriction determines the outcome of the standard setting, which is the special case of homogeneous distances. A Nash equilibrium of the game Γ^δ is also called a locational (Nash) equilibrium (cf. Eiselt and Laporte, 1991, 1993). Thus, $s \in S$ is a *locational equilibrium* if for all $c \in P$ and for all $x \in X$ we have $\pi^c(s^c, s^{-c}) \geq \pi^c(x, s^{-c})$.

Example 4.1. *Consider a cycle graph (X, E) on six nodes, i.e., $X = \{x_1, x_2, \ldots, x_6\}$ and $E = \{x_1x_2, x_2x_3, \ldots, x_6x_1\}$. Let $N = \{i_1, i_2, \ldots, i_{12}\}$ be a set of twelve agents with favorite objects $(\hat{x}^1, \hat{x}^2, \ldots, \hat{x}^{12}) = (x_1, x_1, x_2, x_2, \ldots, x_6, x_6)$. We first assume homogeneous edge lengths, i.e., for all $i \in N$ we have $\delta_e^i = 1$ for any $e \in E$. Together with a set of three players $P = \{c_1, c_2, c_3\}$ this constitutes a game Γ^δ.*

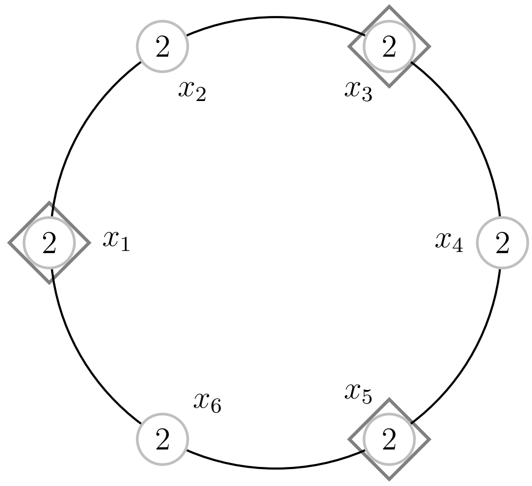


Figure 4.2: Three players on a cycle graph.

The graph (X, E) is illustrated in Figure 4.2. The number within a node indicates the number of agents who have this node as the favorite object. The edge lengths are not represented. Finally, the three squares represent the strategy profile $(s^1, s^2, s^3) = (x_1, x_3, x_5)$. We will keep these conventions in the following figures.
For this game, results of Mavronicolas et al. (2008) imply that the depicted strategy profile s is a locational equilibrium. A player cannot improve by relocating, because

her payoff either remains 4 *(when deviating to a neighbor) or decreases. This result, however, depends on the specific edge lengths. Consider the situation where one of the twelve agents with favorite object on* x_2 *assigns a different length to an edge next to him, such as* $\tilde{\delta}^3_{x_1x_2} = 1 - \epsilon$ *for some* $\epsilon > 0$ *and* $\tilde{\delta}^3_e = 1$ *for all other edges. The perceived distances of the other agents are assumed to stay the same. Then the depicted strategy profile* s *is not a locational equilibrium. The player* $c_3 \in P$ *with strategy* x_3 *now has an incentive to deviate to* x_2 *or* x_4 *because in both cases she would attract four agents instead of only 3.5. Thus, the strategy profile* $s \in S$ *is a locational equilibrium in the game* Γ^δ *but not in the perturbed game* $\Gamma^{\tilde{\delta}}$*. In some sense the profile is not "robust".*

The previous example motivates the following definition:

Definition 4.1 (Robust equilibrium)**.** A strategy profile $s^* \in S$ is a *robust equilibrium* if it is a locational equilibrium for any collection of individual edge lengths. In other words: $s^* \in S$ is a locational equilibrium in Γ^δ for any $\delta = (\delta^i_e)^{i \in N}_{e \in E}$.

Certainly, robustness is a strong requirement. But it is a desirable property for at least two reasons. First, a robust equilibrium is independent of the assumption of homogeneous edge lengths but includes this as a special case. Indeed, a robust equilibrium is also a locational equilibrium in the homogeneous case Γ^δ, where $(\delta^i_e)_{e \in E}$ is the same for all agents $i \in N$. Second, to determine the locational equilibrium one has to specify for each agent her favorite object $\hat{x}^i \in X$ as well as her list of edge lengths $(\delta^i_e)_{e \in E}$ together with a graph (X, E). On the other hand, to determine robust equilibria it is sufficient to know the graph (X, E) and the distribution of favorite objects $(\hat{x}^i)^{i \in N}$. In fact, it is sufficient to have only information about the node-weighted graph that is induced by $(\hat{x}^i)^{i \in N}$, i.e., it is enough to know (X, E, w) where $w_x := |\{i \in N \mid \hat{x}^i = x\}|$ is the number of agents having x as their favorite object. We will interpret an exogenously given node-weighted graph (X, E, w) as a situation of spatial competition.

4.3 Robustness

We will first give a characterization of robustness which applies to test whether locational equilibria are robust. Then, we will turn to properties of robust equilibria, in particular minimal differentiation and efficiency. Finally, we will reconsider the existence of robust equilibria.

4.3.1 Characterization

In this subsection we provide the necessary and sufficient conditions for a strategy profile to be a robust equilibrium. For this purpose we need additional definitions.

Definition 4.2. Let (X, E) be a graph and fix a strategy profile $s \in S$. Furthermore, let $\bar{X} = \bigcup_{c=1}^{p} \{s^c\} \subseteq X$ be the set of occupied nodes in s.

- The *hinterland* $H_x \subseteq X$ of node $x \in \bar{X}$ is the set of nodes that have x on every path to any $x' \in \bar{X}$. In the special case where all players choose the same strategy (i.e., $|\bar{X}| = 1$), say $\bar{X} = \{x\}$, we define $H_x := X$.

- An *unoccupied zone* $Z \subseteq X$ is a component of $(X, E) - \bar{X}$. The set of all unoccupied zones is denoted by $\mathcal{Z}$.

- An unoccupied zone $Y \subseteq X$ is called a *competitive zone* if it is not contained in any hinterland, i.e., $Y \nsubseteq H_x$ for all $x \in \bar{X}$. The set of all competitive zones is $\mathcal{Y}$.

- Two distinct objects $x, x' \in \bar{X}$ are *indirectly neighboring* if there exists a competitive zone to which both nodes are neighboring.

- The *neighboring area* $A_x \subseteq X$ of $x \in \bar{X}$ is the unoccupied zone which would be obtained when removing all players located on x. Formally, that is $A_x = (\bigcup_{Z \in \mathcal{Z}_x} Z \cup \{x\})$, where $\mathcal{Z}_x := \{Z \in \mathcal{Z} \mid Z \text{ neighboring to } x\}$.

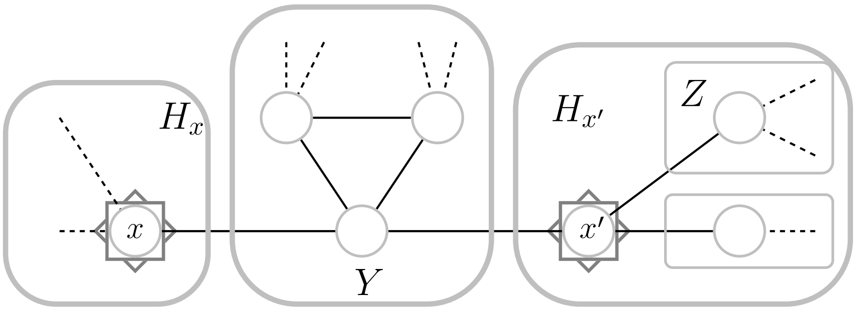


Figure 4.3: Example for definitions: decomposition into competitive zones and hinterlands.

The notions of hinterland and competitive zone go back to Eiselt (1992) who has defined them for the given positions of two players. The hinterland $H_x \subseteq X$ consists of the node itself and possibly several unoccupied zones that are adjacent to $x \in \bar{X}$ but not to any other occupied node in $\bar{X}$. Agents who have their favorite object in H_x must be closer to player(s) on node x than to all other players, since any path, and therefore also the shortest one(s), contain this object. This is different for competitive zones. Players surrounding a competitive zone $Y \in \mathcal{Y}$ compete with indirectly neighboring competitors over the agents who have their most favorite object in Y. The definitions are illustrated in Figure 4.3, where there are two occupied nodes $x, x' \in \bar{X}$, several unoccupied zones, where one of them (Y) is a competitive zone, and another one (Z) belongs to a hinterland. Furthermore, the neighboring area $A_x \subseteq X$ consists of the hinterland H_x and the competitive zone Y, while the neighboring area $A_{x'}$ consists of the other hinterland $H_{x'}$ and the competitive zone Y. Generally, each node either belongs to one hinterland or to one competitive zone. This can be considered as a partition of X into l hinterlands (i.e., $|\bar{X}| = l$) and k competitive zones

$$\Pi(s) = \{H_{x_1}, \dots, H_{x_l}, Y_1, \dots, Y_k\}. \tag{4.1}$$

In fact, because every agent with favorite object in $H_x \subseteq X$ is always closer to a player on the corresponding node x than to any other occupied node, $\frac{w(H_x)}{p_x}$ is the "worst-case payoff" that a player who chooses x receives. Conversely, the maximal payoff of a player who chooses x is restricted by the neighboring area $A_x \subseteq X$, i.e., by $\frac{w(A_x)}{p_x}$. These simple considerations lead to the following key proposition.

Proposition 4.1. Let $s^* \in S$ be a strategy profile on a node-weighted graph (X, E, w) and let $\Pi(s)$ be the corresponding partition as in (4.1). Furthermore, let $\hat{Z} \in \operatorname{argmax}_{Z \in \mathcal{Z}} w(Z)$ be a heaviest unoccupied zone. Then s^* is a robust equilibrium if and only if the following four conditions are satisfied for all $x \in \bar{X}$:

$$\text{(1.)} \qquad \frac{w(H_x)}{p_x} \geq w(\hat{Z})$$

$$\text{(2.)} \qquad \frac{w(H_x)}{p_x} \geq \frac{w(A_{x'})}{p_{x'} + 1} \qquad \forall\, x' \in \bar{X} \setminus \{x\}$$

Furthermore, if $p_x = 1$:

$$\text{(3.)} \qquad w(Y) = 0 \qquad \forall\, Y \in \mathcal{Y}, Y \subseteq A_x$$

$$\text{(4.)} \qquad w(H_x) \geq \frac{w(A_{x'})}{p_{x'}} \qquad \forall\, x' \text{ ind. neighb. to } x.$$

The proof is relegated to the appendix. Proposition 4.1 formalizes the requirements for a strategy profile to be a robust equilibrium. It consists of four straightforward conditions. The first one formalizes that deviations into unoccupied zones are never beneficial. Even if the players only receive their worst case payoff, i.e., the weight of their hinterland, they never gain from relocating into any $Z \in \mathcal{Z}$.[7] Similarly, Condition (2.) captures that deviations to already occupied nodes $x' \in \bar{X}$ are not beneficial. The highest possible payoff a deviating player could get is $\frac{w(A_{x'})}{p_{x'}+1}$.[8] These two previous considerations must be strengthened when considering certain deviations of an isolated player because her node becomes unoccupied then. Again, we distinguish between deviations into a neighboring zone and deviations on occupied nodes, which is reflected by Conditions (3.) and (4.). The main intuition is that for some distance perceptions an isolated player would attract only her hinterland, but by deviating she could receive her former hinterland and, in addition, the weight of some competitive zone (Condition (3.)). By deviating on a neighboring occupied node she can not only share the payoffs of the players on this node, but would also regain some share of her former hinterland (Condition (4.)). For competitive zones neighboring a singly occupied node this means that their weight must be zero. We have already seen an example where this condition is violated. In Example 4.1 there are several singly occupied nodes which are neighboring a non-trivial competitive zone (cf. Figure 4.2).[9] Thus, we can immediately conclude that the given strategy profile is not a robust equilibrium.

The main importance of Proposition 4.1 is that it provides an efficient tool for verifying whether a strategy profile $s \in S$ (which might be a Nash equilibrium for specific edge length, for example) constitutes a robust equilibrium or not. A straightforward algorithm is simply to (i) determine the partition $\Pi(s)$, (ii) compute the weights of the hinterlands and competitive zones, and (iii) check if the four conditions characterizing a robust equilibrium are satisfied. In particular, since the algorithm proceeds in quadratic time, finding a robust equilibrium is as complex as finding a Nash equilibrium for specific edge length. In the remainder of this subsection we will exemplify this for some prominent results from the literature.

[7] This requirement also implies that the weight of unoccupied zones can never be higher than the average payoff of the players, i.e., $w(Z) \leq \frac{n}{p}$ for all $Z \in \mathcal{Z}$.

[8] A simple implication of this requirement is that in robust equilibria the number of players on occupied nodes is roughly proportional to the weights of the hinterlands: $\frac{p_x}{p_{x'}+1} \leq \frac{w(H_x)}{w(H_{x'})} \leq \frac{p_x+1}{p_x}$ for all $x, x' \in \bar{X}$.

[9] We say that a competitive zone Y is trivial if no agent has his favorite object there, i.e., $w(Y) = 0$.

Hotelling's main result for two players on a continuous line is that both cluster on the so-called median. This finding is driven by the fact that both players tend to the center of the line to steal agents from the other player. This is illustrated for a discrete line in Figure 4.4 where we can observe the incentive to increase the hinterland by moving to the discrete analogue of the median.

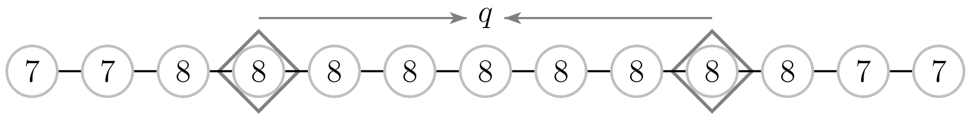


Figure 4.4: A node-weighted line graph with two players.

Definition 4.3 (Median)**.** A *median* of a node-weighted graph (X, E, w) is a node $q \in X$ that balances the node weights, i.e., $w(C_1^q) \leq \frac{w(X)}{2} = \frac{n}{2}$, where $C_1^q \in \mathcal{Z}$ is the heaviest component of $(X, E) - \{q\}$.

In general, a median need not exist. For example, if we consider the complete graph where all weights are equal to one, we have $w(C_1^q) = n - 1 > \frac{n}{2}$. Nevertheless, one can show that if (X, E) is a tree, a median always exists.[10]

The most direct way to extend Hotelling's model to graphs is to consider trees. Although this is only a special case of our set-up, much attention has been devoted to this particular class in literature. Among others, Eiselt and Laporte (1991) examined this setting and they have shown that in the two-player case for homogeneous distances both players will locate on the median of the tree. Thus, they came to the same conclusion as Hotelling did. In fact, this result had already been established by Wendell and McKelvey (1981) in slightly different terms. In their publication the authors show that for homogeneous distances on a tree the median is always a Condorcet winner.[11] Since a Condorcet winner cannot be beaten in majority voting (by definition), choosing the Condorcet winner constitutes a locational equilibrium in the two-player game.

Now, let us apply Proposition 4.1 to test whether the two-player results mentioned in the previous paragraph are robust. If both players locate on the same object, say $q \in X$, there is only one hinterland consisting of all the nodes, i.e.,

[10]Moreover, for trees a node q is a median if and only if $q \in \operatorname{argmin}\left\{ \sum_{y \in X} d(x, y) w_y \;\middle|\; x \in X \right\}$ for all δ (see Goldman, 1971), i.e., a median q is a minimizer of the weighted sum of graph distances for all δ. On general graphs there are multiple conventions for the notion 'median': sometimes it is defined (rather than characterized) as the minimizer of the weighted sum of graph distances.

[11]Later Hansen *et al.* (1986) extended this work.

$\Pi(s) = \{X\}$. Therefore, only Condition (1.) of Proposition 4.1 applies and it simplifies to $\frac{n}{2} \geq w(\hat{Z}) = w(C_1^q)$, which is exactly the definition of the median.[12] Now consider the setting where the players choose different positions, say x and $x' \in X$. Eiselt and Laporte (1991) show that this is a locational equilibrium only if the positions are either neighboring or the competitive zone between them has weight 0 and, furthermore, $\frac{n}{2} = w(C_1^x) = w(C_1^{x'})$ holds. Applying conditions (3.) and (4.) of Proposition 4.1 yields that this is robust, too.
In Eiselt and Laporte (1993) the authors examine the case of three players on a tree. In their main result they distinguish four different cases: (i) type A equilibria (all players cluster on the median $q \in X$), (ii) type B equilibria (two players locate on the median q and one in the heaviest component $C_1^q \in \mathcal{Z}$ on the node that is neighboring to q), (iii) type C equilibria (all three players on different nodes), and (iv) non-existence of equilibria. With the conditions given in Eiselt and Laporte (1993) it is easy to check that type A and type B equilibria are indeed robust. However, type C equilibria generically are not. They are robust only if the hinterland of all players has the same weight because otherwise Condition (4.) of Proposition 4.1 would be violated.

Note that in the previous examples the equilibria are robust only if some kind of minimal differentiation is satisfied and at least some players choose the median q. Therefore these results raise some questions regarding the general form of robust equilibria.

4.3.2 Minimal Differentiation

Minimal differentiation is one of the most controversial results and much attention has been devoted to its implications.[13] In the framework of graphs, we define minimal differentiation as follows.

[12] In fact, this has already been shown for the continuous line, although in very different terms, by the seminal contribution of Black (1948). He proved that for single-peaked preferences on a line the median is always a Condorcet winner. As already mentioned in Section 4.2, single-peaked preferences on a line are is equivalent to our assumption of heterogeneous edge lengths on the line graph.

[13] Some works show that generically it is not satisfied (see, e.g., d'Aspremont *et al.*, 1979; Eaton and Lipsey, 1975; Economides, 1986) but others support it for special cases (see, e.g., de Palma *et al.*, 1985, 1990; Hehenkamp and Wambach, 2010). Similar considerations also apply to minimal differentiation on graphs.

Definition 4.4. A strategy profile $s \in S$ satisfies minimal differentiation if all players locate on the same node, i.e., $s = (x, x, \dots, x)$ for some $x \in X$.

In the previous section there were already examples for robust equilibria satisfying minimal differentiation for two or three players.[14] These cases can be extended to arbitrary numbers of players in a straightforward way. Consider the strategy profile $s := (x, x, \dots, x)$ where all players locate on a node $x \in X$. We then have only one hinterland consisting of all the nodes, i.e., $\Pi(s) = \{X\}$. By using the same arguments as in the two-player case one can see that Conditions (2.), (3.), and (4.) of Proposition 4.1 do not apply and, furthermore, Condition (1.) simplifies to $\frac{n}{p} \geq w(\hat{Z})$, where $\hat{Z}$ is the heaviest unoccupied zone. Thus, we get the following corollary.

Corollary 4.1. Let (X, E, w) be a node-weighted graph and $q \in X$. Furthermore, let $C_1^q \in \mathcal{Z}$ be a heaviest component of $(X, E) - \{q\}$. The strategy profile $s = (q, \dots, q)$ is a robust equilibrium if and only if the weight of any component of $(X, E) - \{q\}$ is not higher than the average payoff, i.e.,

$$w\left(C_1^q\right) \leq \frac{n}{p}.$$

Corollary 4.1 shows that it is easy to construct a robust equilibrium for any number of players. The result is also easy to prove without Proposition 4.1 since for $s = (q, \dots, q)$ every player earns the average payoff $\frac{n}{p}$, while the most beneficial deviation leads to the heaviest unoccupied zone C_1^q. Phrased differently, if the heaviest component of the graph without $q \in X$ is relatively light, then there exists a robust equilibrium where all players locate on the same node. In particular, this also implies that q has to be a median of the graph.
Note that in the robust equilibria discussed so far all players are located on or next to the median. Therefore one might suspect that in any robust equilibrium the median must be occupied (if it exists) and that the players cluster on or around it. The following example is a counter-example to this conjecture.

Example 4.2. *Let (X, E, w) be the weighted line graph depicted in Figure 4.5. Furthermore assume that two players locate on each of the nodes with weight 33. As it is easy to check, this strategy profile is a robust equilibrium. The median, however,*

[14] Definition 4.4 captures minimal differentiation in a strong sense. A weaker version of minimal differentiation would be the requirement that there is no unoccupied node between any pair of occupied nodes or, equivalently, that there is no competitive zone.

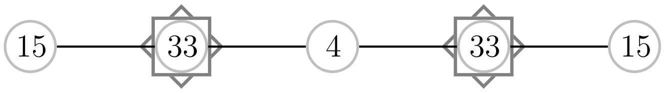


Figure 4.5: A robust equilibrium with no player on the median and without minimal differentiation.

is the node with a weight of four and it belongs to a competitive zone. Thus, neither minimal differentiation is satisfied, nor are players located on the median.
However, consider a reduced game where we remove the two nodes to the right and we remove the two players in this area. In this reduced game, the unique robust equilibrium is that the remaining two players both locate on the node with 33 agents such as in the current strategy profile. Moreover, this node is the median of the reduced graph. A similar observation can be made when reducing the game by removing "the left part".

Example 4.2 shows that in a robust equilibrium it need not be the case that players minimally differentiate on the median. However, it seems that locally, in a kind of reduced game, this is still true. To investigate this issue, let us formally define a reduced game. Given a strategy profile $s \in S$, we define a reduced game for every occupied node $x \in \bar{X}$ by considering the objects and players in the neighboring area $A_x \subseteq X$. Thus, the number of players in the reduced game is p_x and the graph is restricted to $(A_x, E_{|A_x})$. For the payoffs only those agents are considered whose favorite object belongs to the neighboring area A_x such that the node weights of the graph in the reduced game coincide to the node weights of the original game.

Corollary 4.2 (Reduced Games)**.** Suppose $s^* \in S$ is a robust equilibrium for some (X, E, w) and let $x \in \bar{X}$ be an occupied position such that $p_x \geq 2$. Then, x is the median of the subgraph $(A_x, E_{|A_x})$ and $(x, x, \ldots, x)$ is a robust equilibrium satisfying minimal differentiation in the corresponding reduced game.

Proof. Let $x \in \bar{X}$ be an occupied position in $s^* \in S$ with $p_x \geq 2$. Applying Proposition 4.1, Condition (1.) implies $\frac{w(A_x)}{p_x} \geq \frac{w(H_x)}{p_x} \geq w(Z)$ for every unoccupied zone surrounding x. But this is equivalent to the condition of Corollary 4.1, $w(C_1^x) \geq \frac{w(A)}{p_x}$, which shows that the strategy profile $(x, \ldots, x)$ is a robust equilibrium in the reduced game. Moreover, this condition implies that the weight of the heaviest component of $\left(A_x, E_{|A_x}\right) - \{x\}$ is smaller than $\frac{w(A_x)}{2}$ which shows that $x \in \bar{X}$ is the median of the subgraph $(A_x, E_{|A_x})$. $\square$

Corollary 4.2 shows that in any robust equilibrium a local variant of minimal differentiation is satisfied. This finding is fully in line with the “principle of local clustering” conjectured in the seminal work of Eaton and Lipsey (1975). Their principle, however, also contains the aspect that players pair, i.e., do not locate away from other firms. This aspect is also true in robust equilibria since it follows from Condition (3.) of Proposition 4.1 that isolated players do not neighbor a non-trivial competitive zone. This implies that singly occupied nodes must neighbor another occupied node if node weights are strictly positive. Thus, any robust equilibrium can be characterized as a few multiply occupied nodes which are possibly neighbored by some singly occupied nodes. The final question on the extent of differentiation is whether these local clusters can be at a large distance from each other.
In Example 4.2 only a small share of agents favor the object between the occupied positions. In fact, it holds generally that the weight of competitive zones in robust equilibria must be relatively light.

Proposition 4.2 (Competitive zones)**.** Let (X, E, w) be a node-weighted graph. Suppose $s^* \in S$ is a robust equilibrium and let $\mathcal{Y}$ be the set of competitive zones. Then, $\sum_{Y \in \mathcal{Y}} w(Y) \leq \frac{n}{5}$.

The proof can be found in the appendix. By definition, a strategy profile satisfies minimal differentiation only if there is no competitive zone. In this context, Proposition 4.2 can be interpreted as a weaker form of a global minimal differentiation result: competitive zones might exist in equilibrium, but their weight in sum is bounded by $\frac{n}{5}$, i.e., at most 20% of the agents can have their favorite object in some competitive zone.

The requirement of robustness is crucial for each of the results on minimal differentiation. Indeed, it is possible to find (non-robust) locational equilibria which do not satisfy the properties specified by Corollary 4.1, Corollary 4.2, and Proposition 4.2. Whether robustness also leads to stronger results with respect to efficiency is addressed next.

4.3.3 (In-)Efficiency

Traditionally, welfare is measured by aggregating the players’ and the agents’ surplus. However, from the players’ perspective, in our setting (i.e., without considering price competition) any strategy profile yields the same aggregated surplus as we study a constant-sum game. Therefore, efficiency will be discussed from the viewpoint of the agents which are interpreted as consumers in this subsection.[15] The standard result of two firms choosing the median of a line is known to be inefficient since minimal differentiation leads to unnecessarily high distances for the consumers. In his paper, Hotelling complains about this inefficiency:

> "Buyers are confronted everywhere with an excessive sameness [...]" and "[...] competing sellers tend to become too much alike."
>
> (Hotelling, 1929, p. 54)

This result, however, does not simply generalize. Reconsider Example 4.1 where some agents are uniformly distributed along a cycle graph with equal edge lengths. The (non-robust) locational equilibrium depicted in Figure 4.2 is efficient with respect to different criteria. For instance, it minimizes the sum of distances (of each consumer to a closest player) as well as the sum of squared distances, which are the most common cardinal criteria.[16] However, the cardinal approach does not seem to be fully justified in our context as we have individual distance perceptions which need not be comparable across consumers. A well-known ordinal criterion is Pareto efficiency. The locational equilibrium in Example 4.1 satisfies this criterion as well, i.e., there does not exist another strategy profile such that any consumer is at least as well off and at least one consumer is strictly better off (where better off here means that the perceived distance to the closest player becomes shorter). Note that this is a weak requirement which is satisfied by plenty of strategy profiles. The existence of locational equilibria that are efficient therefore raises the question of whether robust equilibria can be efficient as well. Under generic conditions, the answer is no.

Proposition 4.3 (Pareto efficiency)**.** Let (X, E, w) be a node-weighted graph. Suppose that the number of agents n is not divisible by the number of players p and that there are at least p nodes with positive weight $w_x > 0$. Then any robust equilibrium is Pareto dominated (for the consumers).

Proposition 4.3 shows that under mild conditions robust equilibria are not Pareto efficient. This statement of inefficiency with respect to an ordinal criterion precludes

[15] These might be inhabitants that visit a facility or consumers who buy a product. Because we have not specified a second stage like government formation in our model, the discussion of efficiency does not apply to the context of voting.

[16] The sum of squared distances as an efficiency criterion has been used, for example, by Meagher and Zauner (2004) and Król (2012) who find different effects of uncertainty on efficiency.

inefficiency with respect to cardinal criteria as well since no Pareto dominated strategy profile can minimize the sum of (squared) distances. The proof of Proposition 4.3 is relegated to the appendix. Its intuition is simple. Generically, in every robust equilibrium there are two firms that choose the same location, while the consumers would benefit if one of them located at a different position. In fact, as we have a constant-sum game between the players, a social planner could relocate them and provide transfer payments to keep their payoffs constant. Thus, a socially optimal outcome from the consumers' point of view would be possible without changing the payoffs of the players.[17] This shows that, in a much more general form, Hotelling's inefficiency persists when robustness is required.

4.3.4 (Non-)Existence of Robust Equilibria

So far we analyzed properties of robust equilibria without explicitly examining under which conditions they exist. In Subsection 4.3.1 we have shown for small numbers of players on tree graphs that most of the sufficient conditions from the literature indeed induce robust equilibria. Moreover, Corollary 4.1 provides a condition which is sufficient for existence. Intuitively, it is satisfied either if the weight is concentrated on the median or if we have a star-like structure under a more equal weight distribution. Although this condition is necessary and sufficient only for robust equilibria with minimal differentiation, similar considerations also apply in general. Corollary 4.1 is based on Proposition 4.1 which characterizes the underlying strategy profiles of robust equilibria.[18] In particular, Condition (1.) states that the hinterland $H_x \subseteq X$ of every occupied node $x \in \bar{X}$ must be heavy enough to carry p_x players. If this weight is not directly on the node x, then it must be on other nodes in its hinterland. Considering the "arms" in the hinterland, i.e., the components in the graph $(H_x, E_{|H_x}) - \{x\}$, each of them is an unoccupied zone. However, for unoccupied zones the weight is bounded, again by Proposition 4.1 Condition (1.). Thus, in order to be heavy enough, an occupied node $x \in \bar{X}$ must either have sufficiently

[17] However, this result also depends on the abstraction from price competition. If firms do not cluster, i.e., if they have a local monopoly, they might have an incentive to raise prices.

[18] Proposition 4.1 provides the necessary and sufficient conditions for existence in the sense that a robust equilibrium exists if and only if there is a strategy profile that satisfies these conditions. Thus, this result transforms the problem of finding a strategy profile that is a robust equilibrium into finding a strategy profile that satisfies the conditions of Proposition 4.1, but it is not a result on the exogenously given situation of spatial competition, i.e., on the node-weighted graph (X, E, w).

many arms in its hinterland (which are heavy in sum) or it must have a relatively high weight itself. This intuition is formalized in Corollary 4.3.

Corollary 4.3. For some node-weighted graph (X, E, w), let $s^* \in S$ be a robust equilibrium with heaviest unoccupied zone $\hat{Z} \in \mathcal{Z}$ (and $w(\hat{Z}) > 0$). Let $x \in \bar{X}$ be occupied by $0 < p_x < p$ players. Denote by $a_x \in \mathbb{N}$ the number of arms (i.e., the number of components in the hinterland for $(H_x, E_{|H_x}) - \{x\})$ of x. Then

$$\frac{w_x}{w(\hat{Z})} + a_x \geq p_x.$$

Proof. Let $\hat{Z}_x \in \mathcal{Z}$ be the heaviest unoccupied zone in the hinterland of $x \in \bar{X}$. The result then follows from Proposition 4.1 Condition (1.):

$$\begin{aligned} w(H_x) \geq p_x w(\hat{Z}) \quad &\Rightarrow \quad w_x + a_x w(\hat{Z}_x) \geq p_x w(\hat{Z}) \\ &\Rightarrow \quad \frac{w_x}{w(\hat{Z})} + a_x \cdot \frac{w(\hat{Z}_x)}{w(\hat{Z})} \geq p_x \quad \Rightarrow \quad \frac{w_x}{w(\hat{Z})} + a_x \geq p_x \end{aligned}$$

□

Corollary 4.3 shows that in a robust equilibrium the relative weight of an occupied node plus its number of arms must exceed the number of players on it. This result is illustrated in Figure 4.6 with two occupied nodes x and $x' \in \bar{X}$.

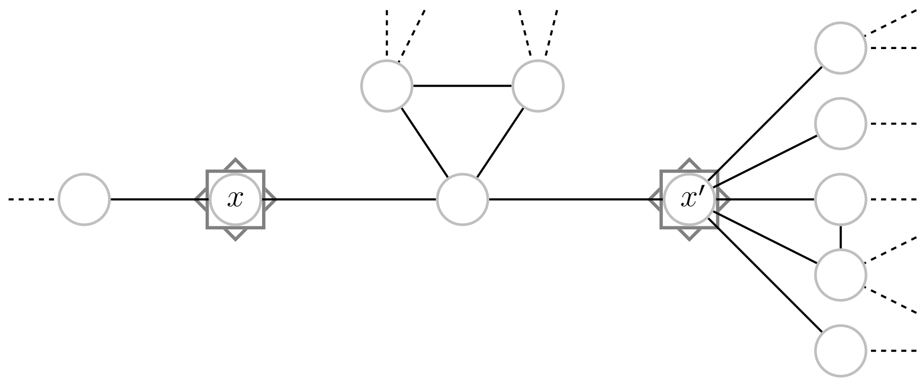


Figure 4.6: Four players on two nodes. If this is a robust equilibrium, then node x must have high weight. This is not necessarily true for x' because it has a high degree (which leads to several arms in its hinterland).

While x has only one arm in its hinterland, x' has four of them. Therefore, for node x we have $\frac{w_x}{w(\hat{Z})} + 1 \geq 2$, which is equivalent to $w_x \geq w(\hat{Z})$, i.e., the weight of the node

must exceed the weight of the heaviest unoccupied zone. Note that this implies an inequality of weights if there are unoccupied zones with many nodes. In contrast to this, x' needs not be as heavy as x, but in order to have four arms it must be a cut vertex and have a degree larger than five. Thus, one interpretation for Corollary 4.3 is that the weight of occupied nodes and their degree can be interpreted as some kind of substitutes: at least one of them has to be high enough in order to carry p_x players in equilibrium.
This gives a requirement for robust equilibria on the level of single nodes. On the graph level this requirement will translate into (a) structural features of the graph and in (b) conditions on the distribution of weights. To assess the weight distribution, we consider the inequality of weights measured by the variance. In our case it is given by $\mathrm{Var}(w) = \sum_{x \in X} \left(w_x - \frac{n}{\xi}\right)^2 = \frac{1}{\xi} \sum_{x \in X} w_x^2 - \frac{n^2}{\xi^2}$. The variance is the quadratic distance from the uniform distribution. In particular, $\mathrm{Var}(w) = 0$ if and only if $w_x = \frac{n}{\xi}$ for all $x \in X$, i.e., if and only if w is uniformly distributed (a special case that is predominantly discussed in the literature). To assess structural requirements of a graph we consider its connectedness which is measured by the number of blocks b (cf. Diestel, 2005). If this number is smaller than the number of players p, then it is still impossible to have Corollary 4.3 trivially satisfied (such as for node x' in Figure 4.6). For these graphs Corollary 4.3 has implications on the weight distribution because there must be an occupied node that is similar to node x in Figure 4.6. As a consequence we have that graphs with a high connectivity (i.e., a relatively small number of blocks) only admit robust equilibria if the weight distribution is far from uniform.

Proposition 4.4. Let (X, E, w) be a node-weighted graph with $\xi > 3p$. Suppose that the number of blocks is smaller than the number of players, i.e., $b < p$. Then there exists some $\nu > 0$ such that $\mathrm{Var}(w) < \nu$ implies that a robust equilibrium does not exist.

The interpretation of this result is as follows: Suppose the graph is not too small ($\xi > 3p$) and the distribution of agents is sufficiently close to the uniform distribution. Then the existence of robust equilibria requires a low connectivity of the underlying graph in terms of that there must be more blocks than players.
Proposition 4.4 obviously applies to all graphs with just one block (i.e., $b = 1$) like grids, for instance. Those graphs are known as *two-connected* and they are characterized by not containing any cut vertex (see, e.g., Diestel, 2005). Indeed, in this case we have $a_x = 0$ for any occupied node $x \in \bar{X}$ (and for any $s \in S$). Thus, if

a two-connected graph is sufficiently large, it always satisfies the requirements of Proposition 4.4 and therefore it does not admit robust equilibria if the weight distribution is too close to uniformity.[19] A particular example of this class of graphs are cycle graphs (as illustrated in Figure 4.2) which have been studied extensively by Mavronicolas *et al.* (2008). Given a uniform distribution of agents (and edge lengths), the authors have shown that there always exists a Nash equilibrium for $\xi > 3p$. However, Proposition 4.4 immediately implies that these equilibria are not robust.

For tree graphs Proposition 4.4 does not apply since trees consist of many blocks. However, for this special class the number of arms is also restricted by some structural property. Since there are no cycles in a tree, each arm in any hinterland leads to a node of degree 1, a so-called *loose end*. Therefore, completely analogous to Proposition 4.4 we can show the following.

Proposition 4.5. Let (X, E) be a node-weighted tree with $\xi > 3p$. Suppose that $e < p$, where e is the number of loose ends. Then there exists some $\nu > 0$ such that $\mathrm{Var}(w) < \nu$ implies that no robust equilibrium exists.

The number of loose ends is a structural feature that is related with the equality of the degree distribution of the graph. The lowest number of loose ends in a tree is attained in the line graph (which has a highly equal degree distribution), while the highest number is attained in the star graph (which has a highly unequal degree distribution). In that sense, Proposition 4.5 shows that the existence of a robust equilibrium on a tree requires either an unequal distribution of weight or an unequal distribution of degree.

To sum it up, robust equilibria certainly exist for structures that are similar to a star graph (Corollary 4.1) or have a highly concentrated distribution of weights. However, for graphs with few cut vertices (i.e., graphs with a low number of blocks) and for tree graphs, robust equilibria can exist only if the weight distribution is not close to uniform. To consider a numerical example for the required inequality: for trees that satisfy the condition $e < p$ of Proposition 4.5 and for cycle graphs (which always satisfy the condition $b < p$ of Proposition 4.4) we can show that there only

[19] The result that two-connected graphs require a sufficient inequality of node weights can also be derived from Proposition 4.2. Since in two-connected graphs any unoccupied node belongs to a competitive zone, Proposition 4.2 implies that we have $w(\bar{X}) \geq \frac{4}{5}n$ in robust equilibria. Thus, there must be at least one node x with $w_x \geq \frac{4}{5}\frac{n}{p}$. That is, to reach an average payoff $\frac{n}{p}$ it is almost enough to attract all agents with favorite object x.

exists a robust equilibrium of three or more players if there is a node $x \in X$ that is at least $\frac{\xi}{p} - 1$ times heavier than some other node $x' \in X$. Thus, if the number of nodes strongly exceeds the number of players in the game (i.e., $\xi \gg p$), those one-dimensional structures do not admit robust equilibria if the weights are uniformly distributed.

4.4 Discussion

Models of spatial competition predominantly deal with three specific applications: (i) firms that strategically locate facilities (e.g., Eiselt and Laporte, 1993), (ii) political candidates who strategically choose a political platform (e.g., de Palma *et al.*, 1990), and (iii) firms that strategically choose a product specification (e.g., Eaton and Lipsey, 1975). In any of the model variations it has been standardly assumed that agents are heterogeneous with respect to their ideal point (i.e., location/policy/product), but homogeneous with respect to the perception of distances. In particular, it must hold that two agents with the same ideal point agree on the ranking of all the other alternatives. In this paper we have introduced a way to relax this strong homogeneity requirement by considering individual distance perceptions. We assess whether model predictions are robust in the sense that they are independent of the perceived distances. Thereby, we confirm robustness of the equilibria found for two and three players on a tree graph by Eiselt and Laporte (1991, 1993). And we find strong support for a conjecture of the "principle of local clustering" articulated by Eaton and Lipsey (1975, p. 46) who further explain that "[t]he principle of minimum differentiation is a special case of the principle of local clustering when the number of firms in the market is restricted to two." In fact, we have shown that all robust equilibria satisfy local clustering in the sense that we have minimal differentiation in each reduced game. An implication of this result is that robust locational choices are not Pareto efficient, which is in line with Hotelling's conjecture. On the other hand, not all results from models of spatial competition are robust with respect to heterogeneous distance perceptions. Especially in graphs without cut vertices the existence of robust equilibria is highly restricted. We illustrate this in an example of uniform distribution of agents along a cycle graph (analyzed by Mavronicolas *et al.*, 2008). Indeed, by discussing general structural conditions for the existence of robust equilibria, we have shown that the existence generically requires a highly unequal distribution of agents. This also raises the question whether there are robust outcomes in the three main applications mentioned at the beginning of this section. For example, Proposition 4.2 implies that at most 20% of the agents may have their favorite object “between” the players. Interestingly, some empirical data on the geographical distribution of inhabitants suggests that the necessary inequality requirements might just be satisfied. According to the United Nations report from 2012 the rate of urbanization in more developed regions was about 78% in 2011 and it is still increasing.[20] In the US it was even higher than 82%, for example. Thus, the popoulation in more developed regions is quite unequally distributed and this suggests that if firms serve only the major cities this might well be a robust equilibrium, despite the inefficiency for consumers who live outside these cities. In the case of product or policy spaces, the exact distribution of consumers is still an open question. But if it should not meet the requirements of robust equilibria, this would lead again to our main motivation that the assumption of homogeneous distances can have a strong impact on the results. In this case, the use of models of spatial competition in these applications has to be reconsidered carefully.

Although we have focused in this paper on just one – yet crucial – aspect of robustness, several other model specifications can be challenged as well. Some of them do not substantially influence our results. For instance, if the assumption that the players do not locate on the edges of the graph was relaxed, then for any robust equilibrium in this more general set-up there exists another one where the players only locate on the vertices and each of them attracts the same set of agents. Moreover, these additional equilibria exist only under very restrictive conditions. Another aspect that could be relaxed is the assumption that ties are broken equally in the case of equal distances. Although it would then be necessary to adapt the formulations of the results, their substance would not change. The reason is that robust equilibria are independent of the perceived distances and, thus, the tie-breaking rule is relevant only if two players locate at the same position.
On the other hand, however, there are also further assumptions which might well play an important role. In particular, we study a simultaneous move game, while models of sequential moves lead to quite different predictions about minimal differentiation (e.g., Loertscher and Muehlheusser, 2011; Prescott and Visscher, 1977), when more than two players are involved.[21] A further major modeling decision is

[20]United Nations, DESA (2012). World urbanization prospects: The 2011 revision.

[21]Also in the literature on sequential location choices the questionable homogeneity assumption is standard. When relaxing this assumption one can find simple three-player examples where the

whether continuous or discrete space is considered. We have contributed to bridging the two corresponding literatures, but it is left for future research to clarify the role of this modeling assumption; for instance, by approximating a continuous space by a discrete space of shrinking steps.

Acknowledgements

We are grateful to Claude d'Aspremont, Chiara Canta, Herbert Dawid, Martina Eikel, Claus-Jochen Haake, Burkhard Hehenkamp, François Maniquet, Gerd Muelheusser, Tom Truyts, Fernando Vega-Redondo, and two anonymous referees for helpful comments and suggestions. Special thanks go to Jan-Philip Gamp for many interesting and inspiring discussions. We also thank session participants at the "Fourth Congress of the Game Theory Society (GAMES 12)" in Istanbul and at the workshop "Models of Influence and Network Theory (MINT 12)" in Paris as well as seminar audiences at CORE (Louvain-la-Neuve) and FUSL (Brussels). Nils Roehl gratefully acknowledges financial support by the German Research Foundation (DFG) within the Collaborative Research Center "On-The-Fly Computing" (SFB 901).

4.5 Appendix: Proofs

Proof of Proposition 4.1

Necessity: Assume $s^* \in S$ is a robust equilibrium and let $\bar{X} = \{x_1, \dots, x_l\}$ be the set of occupied nodes.

First consider the case $p_x = 1$ where $x \in \bar{X}$. Let $c \in P$ be the player with $s^c = x$. We will establish that x is not neighboring a non-trivial competitive zone, i.e., $w(Y) = 0$ for all $Y \in \mathcal{Y}$ neighboring to x. To see this suppose the opposite is true. Fix some arbitrary object $y \in Y$. Because $A_x \subseteq X$ is connected, it is possible to find edge lengths $\left(\bar{\delta}^i_e\right)_{e \in E}$ for all $i \in N$ with $\hat{x}^i \in A_x \setminus Y$ such that $d^i(y) < d^i(x')$ for any occupied position $x' \in \bar{X}$ neighboring to A_x. This implies $d^i(x) < d^i(y)$ because every path in A_x from $\hat{x}^i$ to y passes through x. Furthermore, for all $j \in N$ with favorite object in Y one can choose edge lengths $\left(\bar{\delta}^i\right)_{e \in E}$ such that $d^j(y) < d^j(x'') < d^j(x)$, where $x \neq x'' \in \bar{X}$ is some occupied position also neighboring to Y. Then the payoff of player c is $\pi^c_{\bar{\delta}}(s^*) = w(A_x) - w(Y) < w(A_x) = \pi^c_{\bar{\delta}}(y, s^{*-c})$. Since she can now beneficially

equilibria are not robust.

deviate, s^* is not a robust equilibrium.
Furthermore, if $s^* \in S$ is robust, an isolated player $c \in P$ may never have an incentive to deviate to an indirectly neighboring position $x' \in \bar{X}$. Because the weight of all competitive zones surrounding x equals 0, $\pi^c(s^*) = w\left(H_x(s^*)\right) = w(A_x)$ for all perceptions of distances. Suppose c relocates to x'. Similar as before, it is possible to construct individual distances $\left(\bar{\delta}^i_e\right)_{e \in E}$ for all $i \in N$ such that every agent with favorite object in $A_{x'}$ or A_x strictly prefers x' to any other occupied position, i.e., $\pi^c_{\bar{\delta}}\left(x', s^{*-c}\right) = \frac{w(A_{x'})+w(A_x)}{p_{x'}+1}$. But this implies

$$\pi^c(s^*) = w\left(H_x\right) \geq \underbrace{\frac{w(A_{x'}) + w(H_x)}{p_{x'} + 1} \qquad \forall\ x' \text{ indirectly neighboring to } x.}_{\text{highest possible payoff at occupied and indirectly neighboring nodes}}$$

$$\Leftrightarrow \qquad w(H_x) \geq \frac{w(A_{x'})}{p_{x'}} \qquad \forall\ x' \text{ indirectly neighboring to } x. \tag{4.2}$$

Now let $p_x \geq 1$. Because $s^* \in S$ is supposed to be a robust equilibrium, it is not possible to perturb distances in such a way that a player can increase her payoff. This implies that the payoff she can attain at least has to be greater than the highest possible gain she can reach if she deviates. With similar arguments as in the case $p_x = 1$ this yields

$$\underbrace{\frac{w\left(H_x\right)}{p_x}}_{\text{worst case payoff at } x} \geq \underbrace{w(\hat{Z})}_{\text{best case payoff at unoccupied nodes}},$$

where $\hat{Z} \in Z$ is the heaviest unoccupied zone, and

$$\frac{w\left(H_x\right)}{p_x} \geq \underbrace{\frac{w(A_{x'})}{p_{x'} + 1} \quad \forall\ x' \in \bar{X} \setminus \{x\}.}_{\text{best case payoff at already occupied nodes}} \tag{4.3}$$

If $p_x = 1$, (4.2) already implies (4.3) for indirectly neighboring objects.

Sufficiency: Now assume the requirements from the proposition are satisfied. We have to show that the strategies where p_x players locate at $x \in \bar{X}$ constitute robust equilibria. First consider the case $p_x = 1$, i.e., a singly occupied node. Conditions (3.) and (4.) make sure that the player cannot improve by deviating to a neighboring competitive zone or by deviating to a directly or indirectly neighboring occupied node. Condition (1.) assures that she cannot improve by deviating to any other unoccupied zone and by Condition (2.) she cannot improve by deviating to any other occupied node. Now, let $p_x > 1$. For a player located on $x \in \bar{X}$, Condition (1.) assures that he cannot improve by deviating to any other unoccupied zone and Condition (2.) assures that he cannot improve by deviating to any other occupied node. $\square$

Proof of Proposition 4.2

Let $s^* \in S$ be a robust equilibrium and $\underline{x} \in \bar{X}$ be the position with lowest worst-case payoff, i.e., $\frac{w(H_{\underline{x}})}{p_{\underline{x}}} \leq \frac{w(H_x)}{p_x}$ for all $x \in \bar{X}$. Then Proposition 4.1 Condition (2.) implies

$$\begin{aligned} w(H_{\underline{x}}) &\geq \frac{p_{\underline{x}}}{p_x + 1} w(A_x) = \frac{p_{\underline{x}}}{p_x + 1} \left(w(H_x) + \sum_{Y \in \mathcal{Y},\, Y \subseteq A_x} w(Y) \right) \\ &\geq \frac{p_x}{p_x + 1} w(H_{\underline{x}}) + \frac{p_{\underline{x}}}{p_x + 1} \underbrace{\sum_{Y \in \mathcal{Y},\, Y \subseteq A_x} w(Y)}_{=: w(\mathcal{Y}_x)}. \end{aligned}$$

and, consequently, $w(H_{\underline{x}}) \geq p_{\underline{x}} w(\mathcal{Y}_x)$ for all $x \in \bar{X} \setminus \{\underline{x}\}$, where $w(\mathcal{Y}_x)$ is the aggregated weight of competitive zones surrounding $x \in \bar{X}$.

Case 1: $p_{\underline{x}} = 1$

Here, Proposition 4.1 Condition (3.) implies $w(Y) = 0$ for all $Y \subseteq A_{\underline{x}}$ and, thus, $w(H_{\underline{x}}) \geq p_{\underline{x}} w(\mathcal{Y}_{\underline{x}}) = 0$. Then:

$$\begin{aligned} n = \sum_{x \in \bar{X}} w(H_x) + \sum_{Y \in \mathcal{Y}} w(Y) &\geq \sum_{x \in \bar{X}} p_x \cdot \frac{w(H_{\underline{x}})}{p_{\underline{x}}} + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq \sum_{x \in \bar{X}} p_x \underbrace{w(\mathcal{Y}_x)}_{=0,\text{ if } p_x = 1} + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq 2 \sum_{x \in \bar{X}} w(\mathcal{Y}_x) + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq 2 \left(2 \sum_{Y \in \mathcal{Y}} w(Y) \right) + \sum_{Y \in \mathcal{Y}} w(Y) = 5 \sum_{Y \in \mathcal{Y}} w(Y), \end{aligned}$$

where the last inequality is due to the fact that by definition of competitive zones each $Y \in \mathcal{Y}$ is neighboring to at least two occupied positions.

Case 2: $p_{\underline{x}} \geq 2$

If $p_x = 1$ for all $x \in \bar{X} \setminus \{\underline{x}\}$, again Condition (3.) from Proposition 4.1 implies

$w(\mathcal{Y}_x) = 0$ for all $x \in \bar{X} \setminus \{\underline{x}\}$ and there remains nothing to show. Therefore assume that there exists at least one $x' \in \bar{X} \setminus \{\underline{x}\}$ with $p_{x'} \geq 2$. Again one can exploit Proposition 4.1 Condition (2.):

$$\begin{aligned} w(H_{\underline{x}}) \geq \frac{p_{\underline{x}}}{p_{x'} + 1} w(A_{x'}) \quad &\Leftrightarrow \quad w(H_{\underline{x}}) \geq p_{\underline{x}} w(A_{x'}) - p_{x'} w(H_{\underline{x}}) \\ &\Leftrightarrow \quad w(H_{\underline{x}}) \geq p_{\underline{x}} w(H_{x'}) - p_{x'} w(H_{\underline{x}}) + p_{\underline{x}} w(\mathcal{Y}_{x'}) \end{aligned}$$

and, analogously,

$$w(H_{x'}) \geq \frac{p_{x'}}{p_{\underline{x}} + 1} w(A_{\underline{x}}) \quad \Leftrightarrow \quad w(H_{x'}) \geq p_{x'} w(H_{\underline{x}}) - p_{\underline{x}} w(H_{x'}) + p_{x'} w(\mathcal{Y}_{\underline{x}}).$$

Now the rest of the proof proceeds similarly to Case 1. According to (4.1) we can again decompose the graph in hinterlands and competitive zones and by using $w(H_{\underline{x}}) \geq p_{\underline{x}} w(\mathcal{Y}_x)$ for all $x \in \bar{X} \setminus \{\underline{x}\}$ one gets

$$\begin{aligned} n &= w(H_{\underline{x}}) + w(H_{x'}) + \sum_{x \in \bar{X} \setminus \{\underline{x}, x'\}} w(H_x) + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq p_{\underline{x}} w(H_{x'}) - p_{x'} w(H_{\underline{x}}) + p_{\underline{x}} w(\mathcal{Y}_{x'}) + p_{x'} w(H_{\underline{x}}) - p_{\underline{x}} w(H_{x'}) + p_{x'} w(\mathcal{Y}_{\underline{x}}) \\ &\quad + \sum_{x \in \bar{X} \setminus \{\underline{x}, x'\}} p_x \cdot \frac{w(H_{\underline{x}})}{p_{\underline{x}}} + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq p_{\underline{x}} w(\mathcal{Y}_{x'}) + p_{x'} w(\mathcal{Y}_{\underline{x}}) + \sum_{x \in \bar{X} \setminus \{\underline{x}, x'\}} p_x \underbrace{w(\mathcal{Y}_x)}_{=0, \text{ if } p_x = 1} + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq 2 \sum_{x \in \bar{X}} w(\mathcal{Y}_x) + \sum_{Y \in \mathcal{Y}} w(Y) \\ &\geq 2 \left(2 \sum_{Y \in \mathcal{Y}} w(Y) \right) + \sum_{Y \in \mathcal{Y}} w(Y) = 5 \sum_{Y \in \mathcal{Y}} w(Y) \end{aligned}$$

Again, the last inequality holds because each $Y \in \mathcal{Y}$ is neighboring to at least two occupied positions. □

Proof of Proposition 4.3

Let (X, E, w) be a node-weighted graph and suppose $\frac{n}{p} \notin \mathbb{N}$. We first show that in any robust equilibrium $s^* \in S$ there is at least one node multiply occupied.
Suppose the opposite is true: There is a robust equilibrium $s \in S$ with only singly occupied nodes, i.e., $p_x = 1$ for all $x \in \bar{X}$. Consider two occupied nodes $x, x' \in \bar{X}$ which are directly or indirectly neighboring. Condition (4.) of Proposition 4.1 then

reads $w(H_x) \geq \frac{w(A_{x'})}{1} \geq \frac{w(H_{x'})}{1}$ and $w(H_{x'}) \geq \frac{w(A_x)}{1} \geq \frac{w(H_x)}{1}$ which implies that $w(H_x) = w(H_{x'})$. Since the graph (X, E) is connected, any occupied node $x \in \bar{X}$ is a direct or indirect neighbor of at least one other occupied node and the relation of being a (direct or indirect) neighbor connects all occupied nodes. Therefore, we have $w(H_x) = w(H_{x'})$ for all $x, x' \in \bar{X}$. Moreover, Condition (3.) of Proposition 4.1 implies that all competitive zones must have a weight of zero (because they have a singly occupied node as a neighbor) such that $\sum_{x \in \bar{X}} w(H_x) = n$. Taken together, this yields $w(H_x) = \frac{n}{p}$ for any $x \in \bar{X}$. However, since the weight of each hinterland is determined by a number of agents, we must have $w(H_x) \in \mathbb{N}$, which contradicts our assumption that $\frac{n}{p} \notin \mathbb{N}$.

Thus, in every robust equilibrium there needs to be a multiply occupied node, say $x \in \bar{X}$. Since at least p nodes have a positive weight, there exists an unoccupied node, say $\tilde{x} \in X \setminus \bar{X}$, with $w_{\tilde{x}} > 0$. Changing the strategy of one player with $s^c = x$ to $\tilde{s}^c = \tilde{x}$ is a Pareto improvement because all consumers with $\hat{x}^i = \tilde{x}$ are better off. □

Proof of Proposition 4.4

To show the proposition, assume the opposite is true: that is, assume there exists a robust equilibrium $s^* \in S$. Let $\hat{Z} \in \mathcal{Z}$ be the heaviest unoccupied zone with respect to s^*. Given the requirements of the proposition, we will show that in each robust equilibrium there exists an occupied node which is heavier than $\hat{Z}$. But if the variance becomes small this leads to a contradiction. The proof proceeds in five steps:

Step 1: The ϵ-ν_ϵ-criterion.

Consider the mapping $\|\cdot\|_1 : \mathbb{R}^\xi \longrightarrow \mathbb{R}$ with $\|w\|_1 = \sum_{x \in X} |w_x|$, also known as the Manhattan norm. It is well-know that $\|\cdot\|_1$ is continuous. Thus, for all $\epsilon > 0$ there exists some $\nu_\epsilon > 0$ such that $\|w - w'\|_2 < \nu_\epsilon$ implies $\|w - w'\|_1 < \epsilon$ for all $w, w' \in \mathbb{R}^\xi$, where $\|w - w'\|_2 = \sqrt{\sum_{x \in X}(w_x - w'_x)^2}$ is, as usual, the Euclidean norm. Let $\epsilon := \frac{2p}{5(p+1)} \cdot \frac{n}{\xi}$. Furthermore, in the following let w' be the uniform distribution $w'_x := \frac{n}{\xi}$ for all $x \in X$.[22] Having specified these variables, the ϵ-ν_ϵ-criterion from above implies that there exists some $\nu := \nu_\epsilon^2 > 0$ such that from $\sqrt{\mathrm{Var}(w)} = \|w - w'\|_2 < \sqrt{\nu}$ always $\sum_{x \in X} |w_x - \frac{n}{\xi}| < \epsilon = \frac{2p}{5(p+1)} \cdot \frac{n}{\xi}$ follows. Correspondingly, for the rest of the proof it is assumed that there is given a tupel of node weights $(w_x)_{x \in X}$ (i.e., $w \geq 0$ and $\sum_{x \in X} w_x = n$) with $\mathrm{Var}(w) < \nu$.

[22] Because the fraction $\frac{n}{\xi}$ need not be an integer, the uniform distribution cannot always be induced by allocating n agents to nodes. Still, it is possible to study the node-weighted graph (X, E, w').

Step 2: We establish that $\left|w(\hat{X}) - |\hat{X}|\frac{n}{\xi}\right| < \epsilon$ for all $\hat{X} \subseteq X$.

If $\mathrm{Var}(w) < \nu$, Step 1 implies for all subsets $\hat{X} \subseteq X$,

$$\left|w(\hat{X}) - |\hat{X}|\frac{n}{\xi}\right| = \left|\sum_{x\in\hat{X}}\left(w_x - \frac{n}{\xi}\right)\right| \leq \sum_{x\in\hat{X}}\left|w_x - \frac{n}{\xi}\right| \leq \sum_{x\in X}\left|w_x - \frac{n}{\xi}\right| < \epsilon.$$

Step 3: We establish that $\sum_{x\in\bar{X}} a_x \leq b$.

The main intuition of this step is that all unoccupied zones can be covered by blocks of the graph and we will show that minimal covers of different zones have to be disjoint. Let $Z_x \neq Z'_{x'}$ be two unoccupied zones in the hinterland of x and x', respectively, where $x, x' \in \bar{X}$. Note that $x = x'$ is allowed but, nevertheless, the two zones may not be equal. If it is not possible to find such two zones, $\sum_{x\in\bar{X}} a_x \leq 1$ and there remains nothing to show. According to Section 4.2 let $\mathcal{B}$ be the set of blocks. Obviously $X = \bigcup_{B\in\mathcal{B}} B$ holds. Therefore there exist $\mathcal{B}^{Z_x}, \mathcal{B}^{Z_{x'}} \subseteq \mathcal{B}$ with $Z_x \subseteq \bigcup_{B\in\mathcal{B}^{Z_x}} B$ and $Z_{x'} \subseteq \bigcup_{B\in\mathcal{B}^{Z_{x'}}} B$ such that both sets are minimal with respect to inclusion, i.e., $\hat{\mathcal{B}} \subsetneq \mathcal{B}^{Z_x}$ implies $Z_x \nsubseteq \bigcup_{B\in\hat{\mathcal{B}}} B$ (analogously for $\hat{\mathcal{B}} \subsetneq \mathcal{B}^{Z_{x'}}$). Given the construction of blocks, the two sets $\mathcal{B}^{Z_x}$ and $\mathcal{B}^{Z_{x'}}$ must be disjoint because otherwise there would be a path from Z_x to $Z'_{x'}$ not passing through x and x', which is not possible due to the definition of hinterlands. Thus:

$$\sum_{x\in\bar{X}} a_x = \sum_{x\in\bar{X}} \sum_{Z_x\in\mathcal{Z},\, Z_x\subseteq H_x} 1 \leq \sum_{x\in\bar{X}} \sum_{Z_x\in\mathcal{Z},\, Z_x\subseteq H_x} \left|\mathcal{B}^{Z_x}\right| \leq |\mathcal{B}| = b$$

Step 4: We establish that $w_{x'} \geq w(\hat{Z})$ for some $x' \in \bar{X}$.

As already has been shown in Step 3, the number of hinterlands is bounded by b and, thus, $\sum_{x\in\bar{X}} a_x \leq b < p = \sum_{x\in\bar{X}} p_x$. Therefore there exists some $x' \in \bar{X}$ with $a_{x'} \leq p_{x'} - 1$ and by applying Corollary 4.3 this yields $w_x \geq w(\hat{Z})$. In words: there necessarily exists an occupied node which is heavier than the heaviest unoccupied zone.

Step 5: The final contradiction.

Since the number of hinterlands is smaller than the number of players and because of Proposition 4.2, the average weight of unoccupied zones in hinterlands needs to be relatively high:

$$w(\hat{Z}) \geq \underbrace{\frac{\sum_{x\in\bar{X}} w(H_x) - w(\bar{X})}{\sum_{x\in\bar{X}} a_x}}_{\text{average weight of unoccupied zones in hinterlands}} > \frac{\frac{4}{5}n - w(\bar{X})}{p}$$

Moreover, according to Step 4 this implies that x' must be relatively heavy as well, $w_{x'} > \frac{\frac{4}{5}n - w(\bar{X})}{p}$. But then from Step 2 it follows that

$$\frac{n}{\xi} + \epsilon > \frac{\frac{4}{5}\xi\frac{n}{\xi} - \left(|\bar{X}|\frac{n}{\xi} + \epsilon\right)}{p} \geq \frac{\frac{12}{5}p\frac{n}{\xi} - p\frac{n}{\xi} - \epsilon}{p} = \frac{7n}{5\xi} - \frac{\epsilon}{p}$$

which contradicts $\epsilon = \frac{2p}{5(p+1)} \cdot \frac{n}{\xi}$. Therefore, s^* cannot be a robust equilibrium.

□

References

Berman, O. and Gavious, A. (2007). Location of terror response facilities: A game between state and terrorist. *European Journal of Operational Research*, **177** (2), 1113–1133.

Black, D. (1948). On the rationale of group decision-making. *Journal of Political Economy*, **56** (1), 23–34.

Braid, R. M. (2005). The equilibrium locations of three stores with different selections of differentiated products. *Economics Letters*, **93**, 31–36.

Carpenter, G. S. (1989). Perceptual position and competitive brand strategy in a two-dimensional two-brand market. *Management Science*, **35** (9), 1029–1044.

d'Aspremont, C., Gabszewicz, J. J. and Thisse, J.-F. (1979). On hotelling's "stability in competition". *Econometrica*, **47** (5), 1145–50.

de Palma, A., Ginsburgh, V., Papageorgiou, V. V. and Thisse, J.-F. (1985). The principle of minimum differentiation holds under sufficient heterogeneity. *Econometrica*, **53** (4), 767–781.

—, Hong, G.-S. and Thisse, J.-F. (1990). Equilibria in multi-party competition under uncertainty. *Social Choice and Welfare*, **7** (3), 247–259.

Demange, G. (1982). Single-peaked orders on a tree. *Mathematical Social Sciences*, **3** (4), 389–396.

Denzau, A., Kats, A. and Slutsky, S. (1985). Multi-agent equilibria with market share and ranking objectives. *Social Choice and Welfare*, **2** (2), 95–117.

Diestel, R. (2005). *Graph Theory (Graduate Texts in Mathematics)*. Springer.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of Political Economy*, **65** (2), 135–150.

Eaton, B. C. and Lipsey, R. G. (1975). The principle of minimum differentiation reconsidered: Some new developments in the theory of spatial competition. *The Review of Economic Studies*, **42** (1), 27–49.

Economides, N. (1986). Minimal and maximal product differentiation in hotelling's duopoly. *Economics Letters*, **21** (1), 67–71.

Eiselt, H. A. (1992). Hotelling's duopoly on a tree. *Annals of Operations Research*, **40** (1), 195–207.

— and Laporte, G. (1991). Locational equilibrium of two facilities on a tree. *RAIRO*, **25** (1), 5–18.

— and — (1993). The existence of equilibria in the 3-facility hotelling model in a tree. *Transportation Science*, **27** (1), 39–43.

— and — (1997). Sequential location problems. *European Journal of Operational Research*, **96** (2), 217–231.

—, — and Thisse, J.-F. (1993). Competitive location models: A framework and bibliography. *Transportation*, **27** (21), 44–54.

— and Marianov, V. (eds.) (2011). *Foundations of Location Analysis*, New York: Springer.

Gabszewicz, J. J. and Thisse, J.-F. (1992). Location. In R. Aumann and S. Hart (eds.), *Handbook of Game Theory with Economic Applications*, *Handbook of Game Theory with Economic Applications*, vol. 1, *9*, Elsevier, pp. 281–304.

Godinho, P. and Dias, J. (2010). A two-player competitive discrete location model with simultaneous decisions. *European Journal of Operational Research*, **207** (3), 1419–1432.

Goldman, A. J. (1971). Optimal center location in simple networks. *Transportation Science*, **5** (2), 212–221.

Hansen, P., Thisse, J. F. and Wendell, R. E. (1986). Equivalence of solutions to network location problems. *Mathematics of Opeations Research*, **11**, 672–678.

Hehenkamp, B. and Wambach, A. (2010). Survival at the center–the stability of minimum differentiation. *Journal of Economic Behavior & Organization*, **76** (3), 853–858.

Hotelling, H. (1929). Stability in competition. *The Economic Journal*, **39** (153), 41–57.

Jiang, A. X., Leyton-Brown, K. and Bhat, N. A. (2011). Action-graph games. *Games and Economic Behavior*, **71** (1), 141–173.

Kress, D. and Pesch, E. (2012). Sequential competitive location on networks. *European Journal of Operational Research*, **217** (3), 483–499.

Król, M. (2012). Product differentiation decisions under ambiguous consumer demand and pessimistic expectations. *International Journal of Industrial Organization*, **to appear**.

Labbé, M. and Hakimi, S. L. (1991). Market and locational equilibrium for two competitors. *Operations Research*, **39** (5), 749–756.

Lancaster, K. (1990). The economics of product variety: A survey. *Marketing Science*, **9** (3), 189–206.

Loertscher, S. and Muehlheusser, G. (2011). Sequential location games. *RAND Journal of Economics*, **42** (4), 639–663.

Mavronicolas, M., Monien, B., Papadopoulou, V. G. and Schoppmann, F. (2008). Voronoi games on cycle graphs. In E. Ochmański and J. Tyszkiewicz (eds.), *Proceedings of the 33rd International Symposium on Mathematical Foundations of Computer Science (MFCS'08), LNCS*, vol. 5162.

Meagher, K. J. and Zauner, K. G. (2004). Product differentiation and location decisions under demand uncertainty. *Journal of Economic Theory*, **117** (2), 201–216.

— and — (2005). Location-then-price competition with uncertain consumer tastes. *Economic Theory*, **25** (4), 799–818.

Moulin, H. (1980). On strategy-proofness and single peakedness. *Public Choice*, **35** (4), 437–455.

Mueller, D. C. (2003). *Public Choice III*. Cambridge University Press.

Nilssen, T. (1997). Sequential location when transportation costs are asymmetric. *Economics Letters*, **54** (2), 191–201.

Osborne, M. J. (1995). Spatial models of political competition under plurality rule: A survey of some explanations of the number of candidates and the positions they take. *Canadian Journal of Economics*, **28** (2), 261–301.

Plastria, F. (2001). Static competitivefacility location: An overview of optimisation approaches. *European Journal of Operational Research*, **129** (3), 461–470.

Prescott, E. C. and Visscher, M. (1977). Sequential location among firms with foresight. *Bell Journal of Economics*, **8** (2), 378–393.

ReVelle, C. S. and Eiselt, H. A. (2005). Location analysis: A synthesis and survey. *European Journal of Operational Research*, **165** (1), 1–19.

Rhee, B.-D. (1996). Consumer heterogeneity and strategic quality decisions. *Management Science*, **42** (2), 157–172.

Roemer, J. E. (2001). *Political Competition: Theory and Applications*. Harvard University Press.

Salop, S. C. (1979). Monopolistic competition with outside goods. *The Bell Journal of Economics*, **10** (1), 141–156.

Shiode, S. and Drezner, Z. (2003). A competitive facility location problem on a tree network with stochastic weights. *European Journal of Operational Research*, **149** (1), 47–52.

Wendell, R. E. and McKelvey, R. D. (1981). New perspectives in competitive location theory. *European Journal of Operational Research*, **6** (2), 174–182.