

# Essays on Coalition Formation Theory

## Citation for published version (APA):

Saulle, R. D. (2019). *Essays on Coalition Formation Theory*. Maastricht University.  
<https://doi.org/10.26481/dis.20190905rs>

## Document status and date:

Published: 01/01/2019

## DOI:

[10.26481/dis.20190905rs](https://doi.org/10.26481/dis.20190905rs)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

Essays  
on  
Coalition Formation Theory



**Essays**  
**on**  
**Coalition Formation Theory**

© Riccardo D. Saulle, Maastricht 2019

All rights reserved. No part of this publication may be reproduced, stored in an automated data system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of the author.

This book was typeset by the author using  $\LaTeX$ .

Cover Design obtained by using the package *TikZ*.

Cover Picture by Margherita Preziosi, obtained by using Photoshop.

Printed in Rome (Italy).

# Essays on Coalition Formation Theory

Dissertation

to obtain the degree of Doctor at Maastricht University  
on the authority of the Rector Magnificus Prof. Dr. Rianne M. Letschert,  
in accordance with the decision of the Board of Deans,  
to be defended in public,

on Thursday 5<sup>th</sup> of September 2019 at 14 : 00 hours.

by

Riccardo D. Saulle

**Supervisor:** Prof. Dr. Jean-Jacques Herings

**Co Supervisors:** Dr. Thomas Demuynck, Université Libre de Bruxelles - Bruxelles  
Dr. Christian Seel

**Assessment Committee:** Prof. Dr. Hans Peters (Chair)  
Prof. Dr. Agnieszka Rusinowska, La Sorbonne Université - Paris  
Dr. Ana Mauleon, Université Saint-Louis - Bruxelles  
Dr. Andrés Perea

*"The greatest challenge to any thinker is stating the problem in a way that will allow a solution"*

---

Bertrand Russell





*A mio nonno Domenico*



# Acknowledgments

This thesis owes its existence to the help, support and inspiration of many people.

First, I would like to express my sincere gratitude to my supervisors Jean-Jacques Herings, Christian Seel and Thomas Demuyne. Jean-Jacques is one of the main reason why I came to Maastricht. I am grateful to him for his patience, motivation, and immense knowledge. I thank him for all the advice, comments and assistance in keeping my progress on schedule. I am grateful to Christian for the continuous support to my Ph.D study and related research. I do thank him for the countless hours of conversations we had, for his intellectual generosity and for gave me the structure that I needed (Qualität kommt von quälen). Despite his failure to turn me into the drinker he wished (I should sorry for this, but I did my best...see below), I could not have imagined having a better mentor for my Ph.D study. Just I would please him to destroy the picture of me that he took during the job market.

I am grateful to Thomas for his useful critiques of this research work and for all the time that we spent in his office working on proofs and technicalities. My only regret is that we have been together in Maastricht only for one year. The same point about drinking applies also here.

Second, I would to thank to people that were absolutely crucial in starting my PhD experience, namely Paola Rodano and Filippo Luca Calciano. Paola knows me since I was a bachelor student of Philosophy at La Sapienza University. During her lectures, she always encouraged me to expose my ideas in front of the classroom without being shy, just honest, intellectually honest. Paola helped me to discover in me some dispositions that, either I did not know I had, or I was too insecure to recognize. Filippo was my advisor at the faculty of Economics in Rome. I owe a huge debt to him. My interested for Game Theory was awakened by his passion, his knowledge and his amazing anecdotes about the great game theorists and in particular about the figure of Jean Francois Mertens. Filippo not only helped me for my studies but he also pushed me to get the best from my self and to not to be afraid of new intellectual challenges. Also, he was wise enough to push me to go to Maastricht.

My time in Maastricht has been some of the best of my life. I had the immense luck and privilege to be surrounded by wonderful people. The AE1 department is a great environment full of amazing people that make every day at work truly enjoyable.

I am glad to have shared my office with Henrik "Simple man in yellow pants" Zaunbrecher and I am glad that he was the first friend I made in Maastricht. His sociality and friendliness made up for my many shortcomings. Most important, Henrik was the main organizer of our beer tasting group (about that, I should thanks Sarah for her hospitality and kindness). Many thank to Marcello "Bomber" Negrini, Nickolas "Gentleman"

Gagnon, Caterina, Fortuna e il povero Egidio, and Pierfrancesco for your friendship. All of you have contributed to make memorable not only my stay in Maastricht but also what has been the best day of my life. Thank you for having been a part of it.

I would like to thank Jasmina and Anastasio for my "orange break", almost every day at 16:30 in their office, and for all the chats (and gossip) that we had. I really miss them.

Additionally, I do thank Aline, Anna, Christine, Diogo, Emre, Eveline, Evy, Frauke, Germano, Giang, Hyemi, Kutay, Lina, Max, Maria, Mariana, Matteo, Mehmet, Olga, Peter, Rasmus, Sergio and Toygar for all the memories we shared in the office, parties and drinks.

Finally, I wish to thank my wife with all my love. Despite the distance and all the difficulties, she always found the way to support me. I will never forget it.

Rome, July 2019

*Riccardo D. Saulle*

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Myopic Stable Set for Social Environments</b>	<b>7</b>
2.1	Introduction . . . . .	7
2.2	The Myopic Stable Set . . . . .	9
2.3	General Properties . . . . .	11
2.3.1	Existence and Uniqueness . . . . .	11
2.3.2	Closed Cycles and the Core . . . . .	13
2.3.3	The von Neumann-Morgenstern Stable Set . . . . .	14
2.3.4	Dynamic Stochastic Processes . . . . .	14
2.4	Applications . . . . .	15
2.4.1	Coalition Function Form Games . . . . .	15
2.4.2	Matching . . . . .	18
2.4.3	Network Formation . . . . .	21
2.4.4	Normal-Form Games . . . . .	22
2.5	Conclusion . . . . .	23
2.6	Appendix . . . . .	24
2.7	Supplement . . . . .	33
2.7.1	Examples . . . . .	33
2.7.2	Three-player Simple Games and the vNM Stable Set . . . . .	38
2.7.3	Proper Simple Games and the $\delta$ -Model . . . . .	41
2.7.4	The vNM Stable Set for the Tamura Example of the Knuth Model . . . . .	43
2.7.5	Shapley-Scarf Housing Markets . . . . .	44
2.7.6	Mixed Environments . . . . .	46
<b>3</b>	<b>Bertrand Competition with Asymmetric Costs: A Solution in Pure Strategies</b>	<b>53</b>
3.1	Introduction . . . . .	53
3.2	Model and Solution Concept . . . . .	54
3.3	Bertrand Duopoly . . . . .	55
3.4	Discussion . . . . .	59
<b>4</b>	<b>The Last Will Be the First and The First Last: Segregation in Societies with Positional Externalities</b>	<b>63</b>
4.1	Introduction . . . . .	63

4.1.1	Related Literature . . . . .	65
4.2	Model and Key Definitions . . . . .	66
4.3	Stability . . . . .	68
4.4	Segregation . . . . .	70
4.5	Results for Competitive and Egalitarian Societies . . . . .	74
4.6	Material Efficiency . . . . .	76
4.7	Concluding Remarks . . . . .	76
4.8	Appendix . . . . .	78
<b>References</b>		<b>87</b>
<i>Valorization Addendum</i>		<b>93</b>
<i>Curriculum Vitae</i>		<b>95</b>

# Chapter 1





# Introduction

Coalition formation theory is the collection of disciplines and mathematical tools which deals with the analysis of coalition formation processes. A coalition formation process is a process by which agents get together to coordinate their actions in order to achieve shared goals. Coalition formation is an important and recurrent pattern in human behavior and many social *phenomena* can be studied in terms of coalition formation processes. Examples are: cartel formation, environmental agreements, political party formation and the provision of a local public good.

Coalition formation theory incorporates two different approaches: a *no-binding agreements approach*, focused on the procedural aspects that are involved in the dynamic of the process; a *blocking approach*, focused on the stability of the possible agreements. Certainly, the two approaches differ in how they formalize the interdependence among the agents. The no-binding agreement approach is in the realm of non-cooperative game theory. In this domain, a game is a detailed model of all the moves available to the agents. By contrast, the blocking approach subsumes to cooperative game theory which abstracts away from this level of detail, and describes only the outcomes that result when the agents get together.

As Brandenburger (2007) remarks, the terms non-cooperative and cooperative are perhaps unfortunate. They might suggest that there is no place for cooperation in the former and no place for conflict and competition in the latter. In fact, neither is the case. One part of non-cooperative game theory (the theory of repeated games for example) studies the possibility of cooperation in ongoing relationships. And cooperative game theory embodies not just cooperation among players, but also competition in a particularly strong, unfettered form. Following Brandenburger (2007), I believe that non-cooperative game theory might be better termed procedural game theory and cooperative game theory combinatorial game theory. This would indicate the real distinction between the two branches of the subject, namely that the first specifies various actions and protocols that are available to the agents while the second describes the outcomes that result when the agents come together in different combinations.

The non-cooperative approach to coalition formation theory provides a rich language and develops useful tools to analyze coalition formation processes. One clear advantage of the approach is that it is able to model how specific details of the interaction may impact the final outcome. Furthermore, Nash (1951) argued that non-cooperative games are more fundamental than cooperative games and any cooperative analysis can and should be subsumed under the non-cooperative approach by making communication and bargaining formal moves in a non-cooperative extensive form game. However, one limitation, is that

its predictions may be highly sensitive to details specified by the non-cooperative forms. For this reason it is worth also analyzing more abstract approaches that attempt to obtain conclusions that are independent of such details. The cooperative approach to coalitional games is one such attempt.

Aumann (1989) expresses the idea behind cooperative game theory in the following terms:

"Cooperative theory starts with a formalization of games that abstracts away altogether from procedures and [...] concentrates, instead, on the possibilities for agreement [...]. There are several reasons that explain why cooperative games came to be treated separately. One is that when one does build negotiation and enforcement procedures explicitly into the model, then the results of a non-cooperative analysis depend very strongly on the precise form of the procedures, on the order of making offers and counter-offers and so on. This may be appropriate in voting situations in which precise rules of parliamentary order prevail, where a good strategist can indeed carry the day. But problems of negotiation are usually more amorphous; it is difficult to pin down just what the procedures are. More fundamentally, there is a feeling that procedures are not really all that relevant; that it is the possibilities for coalition forming, promising and threatening that are decisive, rather than whose turn it is to speak [...]. Detail distracts attention from essentials. Some things are seen better from a distance; the Roman camps around Metzada are indiscernible when one is in them, but easily visible from the top of the mountain."

This thesis is a contribution to the blocking approach.

During my PhD at Maastricht University, I developed a firm interest in this field. I was triggered when I realized that literature studies coalition formation process not only modeling different applications using different specific settings, but also using different tools (solution concepts) to make predictions and these tools are rarely connected to each other.

Within this regard, Ray and Vohra (2015b), two main scholars in this field, stated that

"yet as one surveys the landscape of this area of research, the first feature that attracts attention is the fragmented nature of the literature."

I do not fully agree with this statement. On the contrary, I believe that it is possible to argue that the literature is only apparently "fragmented": from one side, different solution concepts share very close ideas and they exhibit similar features; from the other side, it is possible to show that many specific settings can be represented by a more general framework.

This is part of the motivation of my first chapter that is the second chapter of the thesis. In the chapter, we develop a new solution concept (the myopic stable set) that provides a common structure and thereby connects and generalizes different solution concepts in different specific settings.

The myopic stable set is defined for a very general class of social environments (Chwe, 1994). A social environment is a framework to model strategic interaction among individuals and groups and it consists in four ingredients: a finite set of agents, a set of social

states, a collection of preferences profiles and an effectivity correspondence that models the feasible transitions from one social state to another. We only require that the set of social states is a non-empty and compact metric space. Indeed, a social state can encode a lot of different information such as a network structure, an outcome or a matching. Moreover, by specifying the structure of the effectivity correspondence is possible to keep track of the behaviour of all agents involved in the coalition formation process. The generality of the framework allow us to connect the myopic stable set to other existing solution concepts in different specific settings.

The myopic stable set is defined as the smallest closed set of social states satisfying two conditions: for every social state in the set there are no profitable deviations to a social state outside the set; for every social state outside the set there is a sequence of coalitional deviations which approaches the set.

We show that the myopic stable set exists and is non-empty. Under minor continuity conditions, we also demonstrate uniqueness. Furthermore, the myopic stable set is a super-set of the core and of the set of pure strategy Nash equilibria in non-cooperative games. Additionally, the myopic stable set generalizes and unifies various results from more specific environments. In particular, the myopic stable set coincides with the coalition structure core in coalition function form games if the coalition structure core is non-empty; with the set of stable matchings in the standard one-to-one matching model; with the set of pairwise stable networks and closed cycles in models of network formation; and with the set of pure strategy Nash equilibria in finite super-modular games, finite potential games, and aggregative games.

One of the most interesting features of the myopic stable set is that it provides predictions even when the other solution concepts fail. For example, the set of pure strategy Nash equilibria can be empty, the Core can be empty and the von Neumann Morgenstern stable set can fail to exist.

The third chapter of this thesis is an application of the myopic stable set to a game in which the set of pure strategy Nash equilibria is empty. In particular, we examine a Bertrand duopoly with asymmetric marginal costs and a continuous strategy space. Due to the discontinuity of the profit functions, it is not possible to characterize the best reply of the two firms. In fact, the profit of a firm is increasing as long as its price approaches the opponent's price from below. However, when it reaches that price then they will split the market and it experiences a sizable decrease in profit.

In order to provide a pure strategy equilibrium, many microeconomics and industrial organization textbooks contain different arguments based on different *ad hoc* assumptions, such as discretizing the strategy space, or relying on a particular rationing rule that favors the most efficient firm.

In this chapter we prove existence and uniqueness of the Myopic Stable Set for asymmetric Bertrand competition. In particular we consider two versions of the model: one without restriction and one in which predatory prices are not allowed. The uniqueness conditions provided in Chapter 2 do not apply here. Thus, the uniqueness result here is case dependent.

In the last chapter of the thesis we exploit the flexibility of the social environment to study coalition formation processes among heterogeneous agents with rank-dependent preferences. In particular we assume that agents not only care about their absolute wealth, but also about their relative position in the society. This induces a *phenomenon* of positional

externalities. In particular, a positional externality occurs when in this setting a coalitional deviation alters the context within which the social rank is evaluated.

We study two types of societies, competitive societies in which the surplus of a coalition is split according to productivity and egalitarian societies in which coalitions split their surplus equally. We provide a comparative analysis of societies in terms of stability and segregation. A common expectation is that in competitive societies the most advantaged individuals are the most productive and, the other way round, in egalitarian societies the most advantaged individuals are the least productive. However, our prevision is the opposite. Indeed, in a competitive societies, the individuals with the lowest productivity are very attractive as coalition members, since they yield a higher relative payoff to other members of the coalition. Moreover, in egalitarian societies there is no envy among individuals. Thus, individuals prefer to stay with highly productive individuals in order to increase their material payoff.

In particular we prove that for competitive societies, there is always a core stable allocation based on bottom-up segregation, i.e., individuals with adjacent productivities form coalitions and if some individuals are not part of a productive coalition, then these are the most productive individuals. For egalitarian societies, we obtain top-down segregation as the unique core stable prediction. Finally, by proving that for both societies the Core coincides with the myopic stable set, we show that for every non segregated allocation there exists a sequence of coalitionally better replies that leads to a segregated allocation. Thus, not only all the stable allocations are the segregated ones but, also if a society implements a non segregated allocation, there exists a coalition formation process which restores a segregated allocation.

## Chapter 2



# The Myopic Stable Set for Social Environments

## 2.1 Introduction

Models of coalition formation study a widespread and important pattern of human interaction: agents tend to form groups of equally interested individuals, but behave in a non-cooperative way towards outsiders. For example, individuals in a community join forces to provide a local public good, voters create parties to attain their political goals, and firms set up lobby groups to influence policy-makers.

The literature studies coalition formation in many distinct settings, like networks, coalition function form games, and matching models. In this chapter, we focus on a general class of social environments that covers all of these settings and many more. More precisely, we define a social environment on the basis of four components (Chwe, 1994): a finite collection of agents, a set of states, for every agent preferences over the set of states, and an effectivity correspondence that models the feasible transitions from one state to another. We only require that the set of states is a non-empty, compact metric space. As such, in contrast to most settings in the literature, we allow the state space to be infinite.

For such social environments, we define a new solution concept called the *myopic stable set*, abbreviated as MSS. The MSS extends the idea of level-1 farsighted stability by Herings, Mauleon, and Vannetelbosch (2009, 2018) from finite networks to the general class of social environments. The MSS is defined by three conditions, *deterrence of external deviations*, *asymptotic external stability*, and *minimality*. Deterrence of external deviations requires that no coalition benefits by deviating from a state inside the MSS to a state outside the MSS. Asymptotic external stability makes sure that from any state outside the MSS it is possible to get arbitrarily close to a state inside by a sequence of coalitional deviations. Finally, minimality requires that the MSS is minimal with respect to set inclusion.

Our notion of dominance is myopic as agents or coalitions do not predict how their decision to change the current state will lead to further changes by other agents or coalitions. Such a notion is natural in complex social environments where the number of possible states is large and agents have little information about the possible actions other agents may take or the incentives of other agents.

The myopic stable set is closely related to other myopic concepts as the core and the Nash equilibrium. It differs from the ones in the literature that focus on farsightedness (see among others, Chwe, 1994; Herings, Mauleon, and Vannetelbosch, 2009, 2018; Page, Wooders, and Kamat, 2005; Page and Wooders, 2009; Ray and Vohra, 2015), notions which



strengthen farsightedness by using state-dependent expectations (Greenberg, 1990; Xue, 1998; Dutta and Vohra, 2017), and approaches to farsightedness that are inspired by non-cooperative game theory (Herings, Mauleon, and Vannetelbosch, 2004; Dutta, Ghosal, and Ray, 2005). In particular, Greenberg's (1990) optimistic and conservative standards of behavior are popular farsighted notions that, as our MSS, cover a large set of applications.

Our first main result (Theorem 6) shows that every social environment contains at least one non-empty MSS. Moreover, under minor continuity assumptions, we establish uniqueness of the MSS (Theorem 9). The existence and uniqueness results differ from many other popular solution concepts in the literature. For instance, the core and the coalition structure core for coalition function form games can be empty (Bondareva, 1963; Scarf, 1967; Shapley, 1967); the von Neumann-Morgenstern stable set may fail to exist or to be unique (Lucas, 1968, 1992), optimistic and conservative standards of behavior might fail to exist (Greenberg, 1990), and the set of pure strategy Nash equilibria may be empty.

We provide several additional results that provide more insights about the structure of an MSS. For finite state spaces, we fully characterize the MSS as the union of all closed cycles (Theorem 14), i.e., subsets which are closed under coalitional better replies. This result also provides a connection to stochastic processes of coalition and network formation as in Jackson and Watts (2002) and Sawa (2014). For infinite spaces, the union of all closed cycles is found to be a subset of the MSS. This result is helpful in applications and in the comparison to other solution concepts. For instance, any state in the core is a closed cycle and is therefore included in the MSS. As a special case of this result, it follows that the MSS contains the set of pure strategy Nash equilibria in a normal-form game. Next we define a generalization of the weak improvement property (Friedman and Mezzetti, 2001) to social environments and we show that, under weak continuity conditions, the weak improvement property characterizes the collection of social environments for which the MSS coincides with the core. We also show that if the von Neumann-Morgenstern stable set exists, it has a non-empty intersection with any MSS.

We demonstrate the versatility of our results by analyzing the relationship between the MSS and other solution concepts in specific social environments. In particular, we show that the MSS coincides with the coalition structure core for coalition function form games (Kóczy and Lauwers, 2004) whenever the coalition structure core is non-empty; with the set of stable matchings in the one-to-one matching model by Gale and Shapley (1962); with the set of pairwise stable networks and closed cycles in models of network formation (Jackson and Watts, 2002); and with the set of pure strategy Nash equilibria in pseudo-potential games (Dubey, Haimanko, and Zapechelnyuk, 2006) and finite supermodular games (Topkis, 1979 and Milgrom and Roberts, 1990). We fully characterize the MSS for the class of proper simple games.

The structure of the chapter is as follows. Section 2.2 provides the primitives of our general framework of social environments and introduces and motivates the MSS. Section 2.3 establishes existence, non-emptiness, and uniqueness results. Section 2.4 analyzes our solution concept for various settings and relates it to other stability concepts from the literature. Section 2.5 is a conclusion. All proofs can be found in the appendix. Additional material is covered in the online supplement.

## 2.2 The Myopic Stable Set

In this section, we first introduce the concept of a social environment. Next, we introduce the notions of dominance and asymptotic dominance which we then use to define our solution concept, the myopic stable set.

Let  $N$  be a non-empty finite set of individuals. A coalition  $S$  is a subset of  $N$ . The set of non-empty subsets of  $N$  is denoted by  $\mathcal{N}$ . Let  $(X, d)$  be a metric space, where  $X$  denotes our non-empty state space and  $d$  is a metric on  $X$ .<sup>1</sup> Let some state  $x \in X$  be given and let  $\varepsilon \in \mathbb{R}_{++}$ . We define  $B_\varepsilon(x) = \{y \in X \mid d(x, y) < \varepsilon\}$  as the open ball around  $x$  with radius  $\varepsilon$ . The set  $B_\varepsilon(x)$  contains all states in  $X$  that are in an  $\varepsilon$ -neighborhood of  $x$ . For a sequence  $(x^k)_{k \in \mathbb{N}}$  in  $X$ , we write  $x^k \rightarrow x$  if for all  $\varepsilon > 0$ , there is a number  $N \in \mathbb{N}$  such that for all  $k \geq N$ ,  $x^k \in B_\varepsilon(x)$ , i.e., the sequence  $(x^k)_{k \in \mathbb{N}}$  converges to  $x$ .

An *effectivity correspondence*  $E$  associates with each pair of states  $(x, y) \in X \times X$  a, possibly empty, collection of coalitions  $E(x, y) \subseteq \mathcal{N}$ . If  $S \in E(x, y)$ , we say that coalition  $S$  can move from state  $x$  to state  $y$ . If  $E(x, y) = \emptyset$ , then no coalition can move from  $x$  to  $y$ .

Each individual  $i \in N$  has a complete and transitive preference relation  $\geq_i$  over the state space  $X$ . The profile  $(\geq_i)_{i \in N}$  then lists the preferences of all individuals in  $N$ . We denote the asymmetric part of  $\geq_i$  by  $>_i$ , i.e.,  $x >_i y$  if and only if  $x \geq_i y$  and not  $y \geq_i x$ . A social environment is now defined as follows.

**Definition 1** (Social Environment). A *social environment* is a tuple

$$\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$$

consisting of a non-empty, finite set of agents  $N$ , a non-empty, compact metric space  $(X, d)$  of states, an effectivity correspondence  $E$  on  $X$ , and a profile of preference relations  $(\geq_i)_{i \in N}$  over  $X$ .

In Section 2.4, we specify different social environments which correspond to applications such as coalition function form games, one-to-one matching, network formation, and non-cooperative normal-form games.

For a given social environment  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$ , we say that a state  $y \in X$  dominates another state  $x \in X$  if there is a coalition such that (i) it can move from  $x$  to  $y$  and (ii) each of its members strictly prefers  $y$  over  $x$ .

**Definition 2** (Dominance). A state  $y \in X$  *dominates* a state  $x \in X$  under  $E$  if there exists a coalition  $S \in E(x, y)$  such that for every  $i \in S$  it holds that  $y >_i x$ .

An alternative notion is the one of weak dominance. A state  $y \in X$  *weakly dominates* a state  $x \in X$  if there exists a coalition  $S \in E(x, y)$  such that for all  $i \in S$ ,  $y \geq_i x$  and there is at least one  $j \in S$  such that  $y >_j x$ . When we restrict ourselves to settings with a finite state space  $X$ , all theoretical results from this section remain valid when we replace dominance by weak dominance. For settings where  $X$  is infinite, most results remain valid with the exception of Theorem 12 below which provides sufficient conditions for uniqueness.

<sup>1</sup>A metric is a function  $d : X \times X \rightarrow \mathbb{R}_+$  such that (i) for every  $x, y \in X$ :  $d(x, y) = 0$  if and only if  $x = y$ , (ii) for every  $x, y \in X$ :  $d(x, y) = d(y, x)$ , and (iii) for every  $x, y, z \in X$ ,  $d(x, y) \leq d(x, z) + d(z, y)$ .

Let some state  $x \in X$  be given. The subset of  $X$  consisting of all states that dominate  $x$  together with the state  $x$  itself is denoted by  $f(x)$ , so

$$f(x) = \{x\} \cup \{y \in X \mid y \text{ dominates } x \text{ under } E\}.$$

We refer to  $f$  as the *dominance correspondence*. We define the two-fold composition of  $f$  by  $f^2(x) = \{z \in X \mid \exists y \in X : y \in f(x) \text{ and } z \in f(y)\}$ . By induction, we can define the  $k$ -fold iteration  $f^k(x)$  as the subset of  $X$  that contains all states obtained by a composition of dominance correspondences of length  $k \in \mathbb{N}$ , i.e.,  $y \in f^k(x)$  if there is a  $z \in X$  such that  $y \in f(z)$  and  $z \in f^{k-1}(x)$ . Since by definition  $x \in f(x)$ , it holds that, for all  $k, \ell \in \mathbb{N}$ , if  $k \leq \ell$ , then  $f^k(x) \subseteq f^\ell(x)$ . We define the set of all states that can be reached from  $x$  by a finite number of dominations by  $f^{\mathbb{N}}(x)$ , so

$$f^{\mathbb{N}}(x) = \bigcup_{k \in \mathbb{N}} f^k(x).$$

A state  $y$  is said to asymptotically dominate the state  $x$  if starting from  $x$  it is possible to get arbitrarily close to  $y$  in a finite number of dominations.

**Definition 3 (Asymptotic Dominance).** A state  $y \in X$  *asymptotically dominates* a state  $x \in X$  under  $E$  if for all  $\varepsilon > 0$  there exists  $k \in \mathbb{N}$  and a state  $z \in f^k(x)$  such that  $z \in B_\varepsilon(y)$ .

We denote by  $f^\infty(x)$  the set of all states in  $X$  that asymptotically dominate  $x$ . Formally, we have

$$f^\infty(x) = \{y \in X \mid \forall \varepsilon > 0, \exists k \in \mathbb{N}, \exists z \in f^k(x) \text{ such that } z \in B_\varepsilon(y)\}.$$

It is easy to see that the set  $f^\infty(x)$  coincides with the closure of the set  $f^{\mathbb{N}}(x)$ .

We are now ready to define our solution concept, the myopic stable set, abbreviated as MSS.

**Definition 4 (Myopic Stable Set).** Let  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$  be a social environment. The set  $M \subseteq X$  is a *myopic stable set* if it is closed and satisfies the following three conditions:

1. **Deterrence of external deviations:** For all  $x \in M$ ,  $f(x) \subseteq M$ .
2. **Asymptotic external stability:** For all  $x \notin M$ ,  $f^\infty(x) \cap M \neq \emptyset$ .
3. **Minimality:** There is no closed set  $M' \subsetneq M$  that satisfies Conditions 1 and 2.

The solution concept that results when we replace dominance by weak dominance is called *weak dominance MSS*.

Let  $M$  be an MSS. Deterrence of external deviations requires that no coalition of myopic agents can profitably deviate to a state outside  $M$ . Next, asymptotic external stability requires that from any state outside  $M$  it is possible to get arbitrarily close to a state in  $M$  by a finite number of myopic deviations. Observe that the empty set would necessarily violate asymptotic external stability, so any MSS is non-empty.

Although the property of asymptotic external stability resembles a notion of farsightedness, there is an important distinction. In models with farsighted behavior, coalitions deviate from the current state because they expect to profit from a move in some future

period, i.e., after possible subsequent moves by other coalitions. Our definition of asymptotic external stability, however, is myopic in the sense that coalitions deviate only because they see an immediate gain, without anticipating potential future deviations. The condition of deterrence of external deviations also reflects myopic behavior. Myopic agents will not leave the set as this would not result in an immediate gain for all players in a deviating coalition. In contrast, farsighted agents might do so in the hope of gains after subsequent deviations.

Finally, minimality imposes that there is no smaller closed set of states that satisfies deterrence of external deviations and asymptotic external stability. As discussed in Myerson and Weibull (2015), minimality can be interpreted as a requirement of internal stability.

For finite state spaces, the restriction imposed by asymptotic external stability remains unchanged if  $f^\infty$  is replaced by  $f$ . We refer to the property that for all states  $x \notin M$ ,  $f^{\mathbb{N}} \cap M \neq \emptyset$  as *iterated external stability*. For infinite state spaces, the two concepts differ. If one uses iterated external stability instead of asymptotic external stability, we show in the online supplement that an MSS might fail to exist.

An MSS is a minimal set satisfying deterrence of external deviations and asymptotic external stability. Dropping the minimality requirement leads to the concept of a quasi myopic stable set (QMSS) which is useful in the proofs.

**Definition 5** (Quasi Myopic Stable Set). Let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be a social environment. The set  $M \subseteq X$  is a *quasi myopic stable set* if it is closed and satisfies deterrence of external deviations and asymptotic external stability.

## 2.3 General Properties

This section establishes existence of the myopic stable set in general and, under weak additional assumptions, its uniqueness. We also derive some additional structural properties of myopic stable sets that are used in the next section. We provide a brief discussion of the relationship between the MSS and the von Neumann-Morgenstern stable set. Finally, we relate our approach to dynamic models that rely on stochastic processes.

### 2.3.1 Existence and Uniqueness

The first main result is about existence of the myopic stable set.

**Theorem 6** (Existence). Let  $\Gamma$  be a social environment. Then an MSS exists.

Note that the set of all states is a QMSS and the collection of all sets of states that are a QMSS is partially ordered by inclusion. In the proof of Theorem 6, we verify that the partially ordered collection satisfies the conditions for Zorn's lemma and apply it to conclude that there is a minimal QMSS, i.e., an MSS.

Having established existence of an MSS, we now turn to the cardinality of such sets. The following lemma shows that two myopic stable sets cannot be disjoint.

**Lemma 7.** Let  $\Gamma$  be a social environment and let  $M_1$  and  $M_2$  be two myopic stable sets of  $\Gamma$ . Then  $M_1 \cap M_2 \neq \emptyset$ .

Suppose the intersection  $M_1 \cap M_2$  satisfies deterrence of external deviations and asymptotic external stability, i.e., it is a QMSS. Then, by minimality of  $M_1$  and  $M_2$ , the MSS is unique. In the online supplement, we show that for infinite social environments, the intersection  $M_1 \cap M_2$  might not satisfy asymptotic external stability without any additional assumptions. Thus, we impose the following weak continuity assumption.

**Definition 8** (Lower Hemi-continuity of  $f$ ). The dominance correspondence  $f : X \rightarrow X$  is *lower hemi-continuous* if for every sequence  $(x^k)_{k \in \mathbb{N}}$  in  $X$  such that  $x^k \rightarrow x$  and for every  $y \in f(x)$  there is a sequence  $(y^k)_{k \in \mathbb{N}}$  in  $X$  such that for all  $k$ ,  $y^k \in f(x^k)$  and  $y^k \rightarrow y$ .

This continuity assumption allows us to state the following uniqueness result.

**Theorem 9.** Let  $\Gamma$  be a social environment such that the corresponding dominance correspondence  $f$  is lower hemi-continuous. Then  $\Gamma$  has a unique MSS.

The continuity condition of Theorem 9 is trivially satisfied when the state space  $X$  is finite. As such, for all applications with a finite state space, we have uniqueness of the MSS.

The dominance correspondence  $f$  is defined in terms of the individual preference relations  $(\geq_i)_{i \in N}$  and the effectivity correspondence  $E$ . To ease the verification of lower hemi-continuity of  $f$ , we provide sufficient conditions on the primitives of a social environment.

As a first condition, we impose lower hemi-continuity of the effectivity correspondence  $E$ . Towards this end, consider, for every  $S \in \mathcal{N}$ , the correspondence  $G_S : X \rightarrow X$  defined by

$$G_S(x) = \{x\} \cup \{y \in X \mid S \in E(x, y)\}, \quad x \in X,$$

which associates to every state  $x \in X$  the union of  $\{x\}$  and the set of states coalition  $S$  can move to from  $x$ .

**Definition 10** (Lower Hemi-continuity of  $E$ ). The effectivity correspondence  $E$  is *lower hemi-continuous* if for every coalition  $S \in \mathcal{N}$  the correspondence  $G_S : X \rightarrow X$  is lower hemi-continuous, i.e., for every sequence  $(x^k)_{k \in \mathbb{N}}$  in  $X$  such that  $x^k \rightarrow x$  and for every  $y \in G_S(x)$  there is a sequence  $(y^k)_{k \in \mathbb{N}}$  such that  $y^k \in G_S(x^k)$  and  $y^k \rightarrow y$ .

Our second condition is continuity of the preferences.

**Definition 11** (Continuity of Preferences). The preference relation  $\geq_i$  of individual  $i \in N$  is *continuous* if for any two sequences  $(x^k)_{k \in \mathbb{N}}$  and  $(y^k)_{k \in \mathbb{N}}$  in  $X$  with  $x^k \rightarrow x$  and  $y^k \rightarrow y$  and, for every  $k \in \mathbb{N}$ ,  $x^k \geq_i y^k$ , it holds that  $x \geq_i y$ .

The two conditions turn out to be sufficient for lower hemi-continuity of the dominance correspondence, thus by Theorem 9 for uniqueness of the MSS.

**Theorem 12.** Let  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$  be a social environment such that the effectivity correspondence  $E$  is lower hemi-continuous and the preferences  $(\geq_i)_{i \in N}$  are continuous. Then the dominance correspondence  $f$  is lower hemi-continuous and  $\Gamma$  has a unique MSS.

In the online supplement, we construct a social environment which satisfies the conditions of Theorem 12, but has more than one weak dominance MSS.

### 2.3.2 Closed Cycles and the Core

In this subsection, we give two general results about the structure of an MSS. The first result relates the MSS to the union of all closed cycles. The second result characterizes the social environments for which the MSS is equal to the core.

**Definition 13** (Closed Cycle). A *closed cycle* of a social environment  $\Gamma$  is a set  $C \subseteq X$  such that for every  $x \in C$  it holds that  $f^\infty(x) = C$ .

Thus, a closed cycle is a subset of  $X$  which is closed under the asymptotic dominance correspondence  $f^\infty$ . We denote the union of all closed cycles by  $CC$ , so  $CC$  contains all states that are part of some closed cycle. The following result characterizes the MSS for social environments with a finite state space as the union of all closed cycles and shows that this union is a subset of the MSS for social environments with an infinite state space.

**Theorem 14.** Let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be a social environment and let  $M$  be an MSS of  $\Gamma$ . It holds that  $CC \subseteq M$ . If the state space  $X$  is finite, we have  $CC = M$ .

A *sink* is a closed cycle which consists of only one state, i.e.,  $f(x) = \{x\}$ . The union of all sinks is called the core.

**Definition 15** (Core). Let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be a social environment. The *core*  $CO$  of  $\Gamma$  is given by

$$CO = \{x \in X \mid f(x) = \{x\}\}.$$

It is well-known that the core may be empty for some social environments. However, if it is not empty, then it is always contained in the MSS by the observation that a sink is a closed cycle which consists of one state and by virtue of Theorem 14.

**Corollary 16.** Let  $\Gamma$  be a social environment and let  $M$  be an MSS of  $\Gamma$ . Then we have  $CO \subseteq M$ .

The next definition is inspired by the finite analogue for normal-form games as presented in Friedman and Mezzetti (2001).

**Definition 17** (Weak (Finite) Improvement Property). A social environment  $\Gamma$  satisfies the *weak finite improvement property* if for each state  $x \in X$ ,  $f^N(x)$  contains a sink and the *weak improvement property* if for each state  $x \in X$ ,  $f^\infty(x)$  contains a sink.

The following theorem provides a characterization for the MSS in social environments with the weak improvement property.

**Theorem 18.** Let  $\Gamma$  be a social environment with a lower hemi-continuous dominance correspondence  $f$ . An MSS of  $\Gamma$  is equal to the core if and only if  $\Gamma$  satisfies the weak improvement property.

It follows easily from the proof of Theorem 18 that the requirement of lower hemi-continuity of  $f$  in Theorem 18 can be weakened to the requirement that  $CO$  is closed.

### 2.3.3 The von Neumann-Morgenstern Stable Set

The von Neumann-Morgenstern (vNM) stable set provides a solution concept for an environment consisting of a set of states  $X$  and a dominance relation on this set (von Neumann and Morgenstern, 1944).

**Definition 19** (vNM Stable Set). Let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be a social environment. The set  $V \subseteq X$  is a *vNM stable set* if it satisfies the following two conditions:

1. **Internal stability:** For all  $x, y \in V$  such that  $x \neq y$  it holds that  $y \notin f(x)$ .
2. **External stability:** For all  $x \notin V, f(x) \cap V \neq \emptyset$ .

Internal stability requires that no state in the set is dominated by another state in the set. External stability requires that every state outside the set should be dominated by a state in the set.

Our notion of asymptotic external stability has a similar flavor as the vNM notion of external stability. However, the vNM stable set looks at one-step dominations while our notion of asymptotic external stability uses asymptotic dominance, which can be seen as an infinite iteration of one-step dominations. In fact, extending the definition of the vNM stable set by allowing for a finite iteration of one-step dominations, i.e., replacing  $f$  by  $f^N$  in Definition 19, has also been advocated by several authors, see Harsanyi (1974), van Deemen (1991), Page and Wooders (2009), and Herings, Mauleon, and Vannetelbosch (2017).

On the other hand, our notion of deterrence of external deviations is quite different from the vNM notion of internal stability. While we allow that a state in the MSS is dominated by another state in the MSS, this is prohibited in the vNM stable set. Moreover, unlike our concept, in the vNM stable set it is allowed that a state in the set is dominated by a state outside the set.

In terms of predictions, first note that an MSS always exists and is unique under weak continuity assumptions, whereas the vNM stable set may not exist and if it exists may fail to be unique. If the vNM stable set exists, there are a few connections between the MSS and the vNM stable set. First of all, both sets contain the core. Second, the intersection between the vNM stable set and the MSS is non-empty as is stated in the next result.

**Theorem 20.** Let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be a social environment for which a vNM stable set  $V$  exists. If  $M$  is an MSS of  $\Gamma$ , then  $M \cap V \neq \emptyset$ .

By Theorem 14, the MSS contains the union of all closed cycles. If the vNM stable set exists, one can show that it contains at least one state from every closed cycle.

It is easily verified that Theorem 20 remains true if the dominance correspondences  $f^N$  or  $f^\infty$  are used in Definition 19.

### 2.3.4 Dynamic Stochastic Processes

Stochastic approaches have been frequently used in non-cooperative settings like normal-form games. Sawa (2014) presents a general framework which extends such a stochastic analysis to cooperative settings. In each period, one of the coalitions that can make a move is randomly selected and chooses one of its moves at random. The move is carried

out with probability 1 if all members of the coalition are strictly better off. If no member is worse off, but at least one agent is indifferent, the move is carried out with a probability strictly between 0 and 1. Otherwise, the move is not carried out. The resulting dynamic process can be thought of as a weak better-response dynamic. To establish a connection to the MSS, we consider a variation of the process in Sawa (2014) in which a coalition only moves with positive probability if all coalition members are strictly better off.

Consider a social environment  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  such that  $X$  is finite. For states  $x, y \in X$ , let  $Q(x, y)$  denote the transition probability from state  $x$  to state  $y$  and let  $Q$  be the matrix of transition probabilities. We say that  $Q$  is *consistent* with  $f$  if for every  $y \in f(x) \setminus \{x\}$  it holds that  $Q(x, y) > 0$  and for every  $y \notin f(x)$  it holds that  $Q(x, y) = 0$ . In particular, the state  $x$  need not change even if the set  $f(x) \setminus \{x\}$  is non-empty.

The next result presents an equivalence between the MSS and the set of recurrent states of the Markov chain  $(X, Q)$ .

**Theorem 21.** Let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be a social environment with finite state space  $X$ , let  $f$  be the corresponding dominance correspondence, and let  $(X, Q)$  be a Markov chain such that  $Q$  is consistent with  $f$ . Then the MSS of  $\Gamma$  is equal to the set of recurrent states of  $(X, Q)$ .

For social environments with a finite number of states, Theorem 21 gives an equivalence between the set of recurrent states of a dynamic process that selects all better responses with positive probability and the MSS. Suppose we replace our dominance correspondence used in the definition of the MSS by requiring that  $y$  *indifference dominates*  $x$  if there exists a coalition  $S \in E(x, y)$  such that for all  $i \in S$ ,  $y \succeq_i x$ . Then the predictions of the alternative version of MSS are equivalent to the set of recurrent states in Sawa (2014) for the case of a finite state space.

The above results do not readily extend to settings where  $X$  is infinite, as the set of recurrent states might be empty. In the online supplement, we discuss a stochastic approach to infinite environments based on irreducibility of the Markov chain and show that its predictions differ drastically from those of the MSS.

## 2.4 Applications

In this section, we illustrate the generality of our setting and the useful common structure of our results by means of four specific models that have been studied extensively in the literature: coalition function form games, one-to-one matching models, models of network formation, and normal-form games. For each of these settings, we first specify the social environment, i.e., the set of individuals  $N$ , the state space  $(X, d)$ , the effectivity correspondence  $E$ , and the preferences  $(\succeq_i)_{i \in N}$ . Subsequently, we discuss how the results from the previous section can be applied.

### 2.4.1 Coalition Function Form Games

A coalition function form game is defined by a tuple  $(N, v)$ , where  $N$  is the set of players and  $v : 2^N \rightarrow \mathbb{R}$  is a characteristic function that assigns to each coalition  $S \subseteq N$  a number  $v(S) \in \mathbb{R}$ , called the coalitional value of  $S$ , with the usual convention that  $v(\emptyset) = 0$ . A



coalition structure is a partition  $\pi$  of  $N$ . It describes how the grand coalition is divided into various sub-coalitions. The collection of all coalition structures, i.e., the collection of partitions of  $N$ , is denoted by  $\Pi$ .

For coalition function form games, we define  $X$  as the set of coalition structures  $\Pi$  together with all individually rational payoff vectors that can be obtained by allocating the coalitional values among the members of the respective coalitions:

$$X = \{(\pi, u) \in \Pi \times \mathbb{R}^N \mid \forall i \in N : u_i \geq v(\{i\}) \text{ and } \forall S \in \pi : \sum_{i \in S} u_i = v(S)\}.$$

Given a state  $x \in X$ , we denote by  $\pi(x)$  the projection to its first component, i.e., the coalition structure, and by  $u(x)$  the projection to its second component, i.e., the payoff vector, so we can write  $x = (\pi(x), u(x))$ . The restriction of the payoff vector  $u(x)$  to the members in coalition  $S$  is denoted by  $u_S(x)$ . The set  $X$  is non-empty since it always contains the state where  $N$  is partitioned into singletons and each player  $i \in N$  receives the payoff  $v(\{i\})$ .

For  $x, y \in X$ , we define  $d(x, y) = 1_{\{\pi(x) \neq \pi(y)\}} + \|u(x) - u(y)\|_\infty$ , where 1 is the indicator function and  $\|\cdot\|_\infty$  is the infinity norm. It is easily seen that  $(X, d)$  is compact. We define preferences  $\geq_i$  over the state space  $X$  by setting  $x \geq_i y$  if and only if  $u_i(x) \geq u_i(y)$ .

For each ordered pair of states  $(x, y)$ , the effectivity correspondence  $E(x, y)$  specifies which coalitions can change state  $x$  into state  $y$ . As an example that imposes some reasonable structure on the effectivity correspondence, we provide a brief outline of the notion of *coalitional sovereignty* (Konishi and Ray, 2003; Kóczy and Lauwers, 2004; Ray and Vohra, 2014, 2015; Herings, Mauleon, and Vannetelbosch, 2017).

When a coalition of players  $S$ —called the leaving players—decides to leave their old coalitions to create a new group, the state changes to a new state  $y$  characterized by a new coalition structure  $\pi(y)$  and a new payoff vector  $u(y)$ . The collection of coalitions of  $\pi(x)$  that are unaffected by this change is denoted by  $\mathcal{U}(x, S)$  and the set of all players in this group by  $U(x, S)$ . Formally, we have  $\mathcal{U}(x, S) = \{T \in \pi(x) \mid S \cap T = \emptyset\}$  and  $U(x, S) = \cup_{T \in \mathcal{U}(x, S)} T$ .

This notation helps us in defining coalitional sovereignty.

**Definition 22** (Coalitional sovereignty). An effectivity correspondence  $E$  satisfies *coalitional sovereignty* if the following two conditions hold:

- (1) *Non-interference*: For every  $x, y \in X$ , if  $S \in E(x, y)$  and  $T \in \mathcal{U}(x, S)$ , then  $S \in \pi(y)$ ,  $T \in \pi(y)$ , and  $u_T(x) = u_T(y)$ .
- (2) *Full support*: For every  $x \in X$ , for every  $S \in \mathcal{N}$ , and for every  $u \in \mathbb{R}^S$  such that for all  $i \in S : u_i \geq v(\{i\})$  and  $\sum_{i \in S} u_i = v(S)$ , there is a state  $y \in X$  such that  $S \in E(x, y)$  and  $u_S(y) = u$ .

Intuitively, non-interference requires that if a coalition  $S$  induces a change from a state  $x$  to a state  $y$ , then the unaffected coalitions in  $\mathcal{U}(x, S)$  are still part of the new coalition structure  $\pi(y)$  and every unaffected player  $i \in U(x, S)$  keeps his old payoff, i.e.,  $u_i(x) = u_i(y)$ . Full support requires that every coalition  $S$  has the opportunity to move to a new state where it has the freedom to redistribute its worth  $v(S)$  at will.

Coalitional sovereignty does not fully specify the effectivity correspondence  $E$ . In particular, it does not specify the payoffs and coalition structure of players that are neither part of the leaving coalition  $S$  nor part of the unaffected players  $U(x, S)$ , i.e., players in

the set  $N \setminus (S \cup U(x, S))$ . We call these players residual players. Indeed, one of the more controversial issues is to what extent the leaving players have the power to influence the coalition structure and payoffs of these residual players; see Shubik (1962), Hart and Kurz (1983), Konishi and Ray (2003), and Ray and Vohra (2014) for related discussions and alternative viewpoints. One frequently used specification is the  $\gamma$ -model (Hart and Kurz, 1983). The  $\gamma$ -model prescribes that the residual players are divided into singletons. This assumption is justified by the viewpoint that a coalition is only maintained if there is unanimous agreement among its members. In our setting, the  $\gamma$ -model imposes the following restriction on the effectivity correspondence.

**Definition 23** ( $\gamma$ -model). The effectivity correspondence  $E$  is induced by the  $\gamma$ -model if it satisfies coalitional sovereignty and

- (3) For every  $x, y \in X$ , for every  $S \in E(x, y)$ , if  $i \in N \setminus (S \cup U(x, S))$ , then  $\{i\} \in \pi(y)$ .

The  $\gamma$ -model associates a unique social environment to each coalition function form game.

One of the most prominent set-valued solution concepts for coalition function form games is the coalition structure core.

**Definition 24** (Coalition Structure Core). Let  $(N, v)$  be a coalition function form game and  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$  be the social environment induced by the  $\gamma$ -model. The *coalition structure core* of  $(N, v)$  is the set of states  $x \in X$  such that, for every coalition  $S \in \mathcal{N}$ ,  $\sum_{i \in S} u_i(x) \geq v(S)$ .

In words, the coalition structure core gives to the members of each coalition at least the payoff they can obtain by forming that coalition.

Kóczy and Lauwers (2004) define the coalition structure core to be accessible if from any initial state there is a finite sequence of states ending with an element of the coalition structure core and each element in that sequence outsider independently dominates the previous element. Accessibility of the coalition structure core thus corresponds to iterated external stability of the coalition structure core with respect to outsider independent domination. The notion of outsider independent domination differs from our notion of a myopic improvement in the  $\gamma$ -model in two ways. First, it gives complete freedom to the treatment of residual players after a move has taken place. Second, improvements for the members of the coalition that moves are not necessarily strict improvements. In the on-line supplement, we give an example where under the requirement of strict improvements of all members involved in a move, as in our dominance correspondence  $f$ , the coalition structure core does not satisfy iterated external stability. However, the definition of an MSS uses asymptotic external stability rather than iterated external stability.

By Theorem 6, there exists at least one MSS. Theorem 25 shows the MSS to be unique and to coincide with the coalition structure core whenever the coalition structure core is non-empty.

**Theorem 25.** Let  $(N, v)$  be a coalition function form game and  $\Gamma$  the social environment induced by the  $\gamma$ -model. Then  $\Gamma$  has a unique MSS. If the coalition structure core  $Y$  of  $(N, v)$  is non-empty, then the MSS of  $\Gamma$  is equal to  $Y$ .

Theorem 25 covers coalition function form games with a non-empty coalition structure core. We now study the predictions of MSS when the coalition structure core is empty. To do so, we analyze the class of proper simple games which play an important role in the analysis of political institutions. These are coalition function form games such that  $v(N) = 1$ , for every  $S \subseteq N$  it holds that  $v(S) \in \{0, 1\}$ , and  $v(S) = 1$  implies  $v(N \setminus S) = 0$ . Moreover, for  $S, T \subseteq N$  such that  $S \subseteq T$ ,  $v(S) = 1$  implies  $v(T) = 1$ . The collection of winning coalitions is denoted by  $\mathcal{W} = \{S \subseteq N \mid v(S) = 1\}$ . A veto player is a player that belongs to every winning coalition and the set of veto players is given by  $S^* = \bigcap_{S \in \mathcal{W}} S$ .

It is well-known that the coalition structure core of a proper simple game is non-empty if and only if there are veto players, so  $S^* \neq \emptyset$ . In that case the coalition structure core, and by Theorem 25 the MSS, is given by the states  $x \in X$  such that  $\pi(x)$  has an element containing  $S^*$  and  $\sum_{i \in S^*} u_i(x) = 1$ .

Let us now consider proper simple games with an empty core. Define the subset  $F(X)$  of  $X$  as the set of states such that its partition contains a winning coalition different from the grand coalition and all other elements of the partition are singletons:

$$F(X) = \{x \in X \mid \pi(x) \cap (\mathcal{W} \setminus \{N\}) \neq \emptyset \text{ and } \forall S \in \pi(x) \setminus \mathcal{W}, |S| = 1\}.$$

If  $x \in F(X)$ , then the payoff vector  $u(x)$  distributes one unit of surplus to the players, each player has a non-negative payoff, and at least one player has a payoff of zero. The payoff vector therefore belongs to a facet of the unit simplex.

**Theorem 26.** Let  $(N, v)$  be a coalition function form game such that  $v$  is a proper simple game with an empty core and let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be the social environment induced by the  $\gamma$ -model. Then the MSS of  $\Gamma$  is unique and equal to  $F(X)$ .

Each state in the MSS of the  $\gamma$ -model for simple games with an empty core contains a winning coalition that is a proper subset of the grand coalition. All other coalitions fall apart into singletons. In the online supplement, we consider the  $\delta$ -model as suggested by Hart and Kurz (1983) where residual players remain together. In payoff terms, the predictions of the  $\gamma$ -model and the  $\delta$ -model are shown to coincide, though in the  $\delta$ -model the non-winning coalitions need not be singletons. In the online supplement, we also compare the predictions of the MSS with the vNM stable set for three-player simple games with either one, two, or three winning two-player coalitions. In the first case, both concepts coincide. In the second case, there are two vNM stable sets each with a continuum of states and containing the finite MSS as a proper subset. In the third case, there are four vNM stable sets and their union contains the MSS as a proper subset.

## 2.4.2 Matching

As a second application, we consider the one-to-one matching model of Gale and Shapley (1962) and the one-to-one matching model of Knuth (1976). In the online supplement, we also discuss the housing matching model of Shapley and Scarf (1974).

Gale and Shapley (1962) introduced a one-to-one matching model for a finite set  $N$  of individuals, partitioned in the two exhaustive subgroups, men  $M$  and women  $W$ . The model can be described by a tuple  $(M, W, (P_m)_{m \in M}, (P_w)_{w \in W})$  of individuals and their preference relations. A matching is a function  $\mu : M \cup W \rightarrow M \cup W$  satisfying the following properties:

1. For every man  $m \in M$ ,  $\mu(m) \in W \cup \{m\}$ .
2. For every women  $w \in W$ ,  $\mu(w) \in M \cup \{w\}$ .
3. For all men  $m \in M$  and women  $w \in W$ ,  $\mu(m) = w$  if and only if  $\mu(w) = m$ .

In this setting, our state space  $X$  consists of all possible matchings  $\mu$ . Since  $X$  is finite, we can endow it with the discrete metric  $d(\mu, \mu') = 1_{\{\mu \neq \mu'\}}$ . Each man  $m \in M$  has a complete and transitive strict preference relation  $P_m$  over the set  $W \cup \{m\}$  and each woman  $w \in W$  has a complete and transitive strict preference relation  $P_w$  over the set  $M \cup \{w\}$ . The preferences of the individuals  $(\succeq_i)_{i \in M \cup W}$  over the set  $X$  are induced by their preferences over their match, i.e., for all  $m \in M$  it holds that  $\mu \succ_m \mu'$  if and only if  $\mu(m) P_m \mu'(m)$  and for all  $w \in W$  it holds that  $\mu \succ_w \mu'$  if and only if  $\mu(w) P_w \mu'(w)$ .

The formulation of the effectivity correspondence allows us to study the consequences of different hypothesis on the matching process. We introduce two common assumptions from the literature on matching. First, every man or woman is allowed to break the link with the current partner, in which case this man or woman and the former partner become single:

- (1) For all  $i \in N$  and  $\mu \in X$  with  $\mu(i) \neq i$ , we have  $\{i\} \in E(\mu, \mu')$  where  $\mu' \in X$  is such that
  - (i)  $\mu'(i) = i$ ,
  - (ii)  $\mu'(\mu(i)) = \mu(i)$ ,
  - (iii) for every  $j \in N \setminus \{i, \mu(i)\}$ ,  $\mu'(j) = \mu(j)$ .

The second assumption is that any man and woman that are currently not matched to each other can deviate by creating a link and thereby leaving their former partners single:

- (2) For all  $m' \in M$ ,  $w' \in W$ , and  $\mu \in X$  with  $\mu(m') \neq w'$ , we have that  $\{m', w'\} \in E(\mu, \mu')$ , where  $\mu' \in X$  is such that
  - (i)  $\mu'(m') = w'$ ,
  - (ii)  $\mu(m') \in W$  implies  $\mu'(\mu(m')) = \mu(m')$ ,
  - (iii)  $\mu(w') \in M$  implies  $\mu'(\mu(w')) = \mu(w')$ ,
  - (iv) for every  $j \in N \setminus \{m', w', \mu(m'), \mu(w')\}$ ,  $\mu'(j) = \mu(j)$ .

Observe that these two conditions are in line with the  $\gamma$ -model of coalitional sovereignty. This completes the description of the effectivity correspondence and thereby of the social environment of the model by Gale and Shapley (1962).

Let  $(M, W, (P_m)_{m \in M}, (P_w)_{w \in W})$  be a matching problem. A matching  $\mu$  is said to be stable if for every  $i \in M \cup W$  it does not hold that  $i P_i \mu(i)$  and if for every pair  $(m, w) \in M \times W$  it does not hold that  $w P_m \mu(m)$  and  $m P_w \mu(w)$ . It can easily be shown that a matching is stable if and only if it is in the core of the social environment  $\Gamma$ .

In their seminal contribution, Gale and Shapley (1962) showed the existence of a stable matching. The following result of Roth and Vande Vate (1990) is helpful in determining the relation between the set of stable matchings and the MSS.

**Theorem 27.** (Roth and Vande Vate, 1990) For every matching  $\mu \in X$  there is a stable matching  $\mu'$  such that  $\mu' \in f(\mu)$ .

Since the set of states is finite in this application, it holds that  $f^{\mathbb{N}}(\mu) = f^{\infty}(\mu)$ . Recalling Definition 17, the result of Roth and Vande Vate (1990) means that  $\Gamma$  satisfies the weak improvement property. For finite settings,  $f$  is always lower hemi-continuous. Thus, by Theorem 18, the MSS of the social environment induced by the one-to-one matching model coincides with the set of stable matchings, which is the statement of the following corollary.

**Corollary 28.** Let  $(M, W, (P_m)_{m \in M}, (P_w)_{w \in W})$  be a matching problem and let  $\Gamma$  be the induced social environment. The MSS of  $\Gamma$  is unique and equal to the set of stable matchings.

An alternative one-to-one matching model is due to Knuth (1976). This model differs from the model of Gale and Shapley (1962) in that no individual is allowed to be single. Therefore, it requires the number of men to be equal to the number of women. If a blocking pair forms, the deserted partners are matched together. The primitives of the matching model are given by a tuple  $(M, W, (P_m)_{m \in M}, (P_w)_{w \in W})$  with  $|M| = |W|$ . A matching is a function  $\mu : M \cup W \rightarrow M \cup W$  satisfying the following properties:

1. For every man  $m \in M$ ,  $\mu(m) \in W$ .
2. For every women  $w \in W$ ,  $\mu(w) \in M$ .
3. For all men  $m \in M$  and women  $w \in W$ ,  $\mu(m) = w$  if and only if  $\mu(w) = m$ .

The state space  $X$  consists of all matchings  $\mu$  satisfying the above three properties and is endowed with the discrete metric  $d(\mu, \mu') = 1_{\{\mu \neq \mu'\}}$ . The preferences of the individuals  $(\succeq_i)_{i \in M \cup W}$  over the set  $X$  are induced by their preferences over their match, i.e., for all  $m \in M$  it holds that  $\mu \succ_m \mu'$  if and only if  $\mu(m) P_m \mu'(m)$  and for all  $w \in W$  it holds that  $\mu \succ_w \mu'$  if and only if  $\mu(w) P_w \mu'(w)$ . If a man and woman create a new link, the effectivity correspondence also requires a link between their deserted partners. Formally, for all  $m' \in M$ ,  $w' \in W$ , and  $\mu \in X$  with  $\mu(m') \neq w'$ , we have that  $\{m', w'\} \in E(\mu, \mu')$ , where  $\mu' \in X$  is such that

1.  $\mu'(m') = w'$ ,
2.  $\mu'(\mu(w')) = \mu(m')$ ,
3. for every  $j \in N \setminus \{m', w', \mu(m'), \mu(w')\}$ ,  $\mu'(j) = \mu(j)$ .

This completes the definition of the effectivity correspondence and thereby of the social environment.

The core of the social environment induced by the Knuth (1976) model is non-empty. Moreover, as shown by Tamura (1993), when there are at least four women, there are preferences and a matching  $\mu \in X$  such that  $f^{\infty}(\mu)$  does not contain a stable matching. In these cases, the MSS contains matchings outside the core and can thus be rather large. We present such an example in the online supplement and also compare the MSS to the vNM stable set for that example.

To obtain an intuition which states outside the core are part of an MSS, recall that, by Theorem 14, the MSS coincides with the union of all closed cycles. Thus, if the MSS contains states outside the core, these states are part of a closed cycle with more than one element. In such a cycle, agents myopically form new matches and eventually come back to the initial match. These additional states are included in the MSS due to two restrictions on the agents. First, agents are myopic and thus only consider deviations which result in an immediate gain. Second, agents are additionally restricted by the effectivity correspondence which only allows for pairwise deviations.

### 2.4.3 Network Formation

As a third application, we look at the model of network formation by Jackson and Wolinsky (1996). A network is given by a tuple  $g = (N, \mathcal{E})$ , where the nodes  $N$  are the players of the network and  $\mathcal{E}$  is the set of undirected edges of the network. An undirected edge is represented as a set of two distinct players. Two players  $i, j \in N$  are linked in  $g$  if and only if  $\{i, j\} \in \mathcal{E}$ . We abuse notation and write  $ij \in g$  if  $i$  and  $j$  are linked in the network  $g$ . The set of all networks with node set  $N$  is denoted by  $\mathbb{G}$ . A value function for player  $i$  is a function  $v_i : \mathbb{G} \rightarrow \mathbb{R}$  that associates payoffs for player  $i$  for each network in  $\mathbb{G}$ . A network problem is thus given by  $(N, \mathbb{G}, (v_i)_{i \in N})$ .

We identify  $X$  with the set  $\mathbb{G}$  of all possible networks on  $N$  and endow it with the discrete metric  $d(g, g') = 1_{\{g \neq g'\}}$ . Every player  $i \in N$  has a preference relation  $\succeq_i$  over the set  $X$  of all possible networks defined by  $g \succeq_i g'$  if and only if  $v_i(g) \geq v_i(g')$ . Let  $g + ij$  be the network obtained from network  $g$  by adding the link  $ij$  to  $g$  and let  $g - ij$  be the network obtained by deleting link  $ij$  from  $g$ .

We follow Jackson and Wolinsky (1996) by considering deviations by coalitions of size one or two and by assuming link deletion to be one-sided and link addition to be two-sided. One-sided link deletion allows every player to delete one of his links:

- (1) For all players  $i \in N$ , all networks  $g \in X$ , and all links  $ij \in g$ ,  $\{i\} \in E(g, g - ij)$ .

Two-sided link addition allows any two players that are currently not linked to change the network by forming a link between themselves:

- (2) For all players  $i, j \in N$ , all networks  $g \in X$  with  $ij \notin g$ , we have  $\{i, j\} \in E(g, g + ij)$ .

This completes the description of our social environment for the network formation model. It is straightforward to adjust the effectivity correspondence to incorporate models of network formation where more than one link at a time can be changed by coalitions of arbitrary size (Dutta and Mutuswami, 1997; Jackson and van den Nouweland, 2005) or where link formation is one-sided (Bala and Goyal, 2000) into our framework. We refer to Page and Wooders (2009) for a more extensive discussion of alternative rules of network formation.

A network  $g$  is said to be pairwise stable (Jackson and Wolinsky, 1996) if for every  $ij \in g$  it holds that  $v_i(g - ij) \leq v_i(g)$  and  $v_j(g - ij) \leq v_j(g)$  and for every  $ij \notin g$  it holds that  $v_i(g + ij) > v_i(g)$  implies  $v_j(g + ij) \leq v_j(g)$ .<sup>2</sup> It is not hard to show that a network

<sup>2</sup>Pairwise stability as defined in Section 2 of Jackson and Wolinsky (1996) is somewhat stronger and also requires that there is no  $ij \notin g$  such that  $v_i(g + ij) > v_i(g)$  and  $v_j(g + ij) = v_j(g)$ . The weaker notion used here is discussed as an alternative in Section 5 of Jackson and Wolinsky (1996) and is also widely used in the literature. For generic network problems, there are no indifferences, so the two definitions are equivalent.

is pairwise stable if and only if it is in the core of the social environment  $\Gamma$  as defined in Definition 15.

Corollary 16 shows that any pairwise stable network is in the myopic stable set. However, it is not necessarily the case that the MSS only contains the pairwise stable networks.

Consider the binary relation  $R$  on  $X$  defined by  $gRg'$  if  $g \in f(g')$ , i.e.,  $g$  can be reached from  $g'$  by a finite number of dominations. Let  $I$  be the symmetric part of  $R$ , i.e.,  $gIg'$  if and only if  $gRg'$  and  $g'Rg$ . Consider the set of equivalence classes  $\mathbb{E}$  induced by  $I$ . Denote the equivalence class of network  $g$  by  $[g]$ , i.e.,  $g' \in [g]$  if and only if  $g'Ig$ . For two distinct equivalence classes  $[g]$  and  $[g']$  write  $[g]P[g']$  if  $gRg'$ . It is easy to see that  $[g]P[g']$  if and only if  $gRg'$  and not  $gRg'$ .

Let  $V$  be the collection of maximal elements of  $(\mathbb{E}, P)$ , i.e.,  $[g] \in V$  if there is no  $[g'] \in \mathbb{E}$  such that  $[g']P[g]$ . Since an element of  $V$  simply represents a closed cycle as defined in Definition 13, the following result follows from Theorem 14.

**Corollary 29.** Let  $(N, \mathbb{G}, (v_i)_{i \in N})$  be a network problem and let  $\Gamma$  be the induced social environment. A network  $g$  belongs to the unique MSS  $M$  of  $\Gamma$  if and only if the equivalence class  $[g]$  belongs to  $V$ , i.e.,  $M = \{g \in X \mid [g] \in V\}$ .

Herings, Mauleon, and Vannetelbosch (2009) define the pairwise myopically stable sets for network problems using pairwise stability as defined in Section 2 of Jackson and Wolinsky (1996). It is not hard to see that the MSS for social environments  $\Gamma$  coincides with the pairwise myopically stable set for generic network problems. For such network problems, Corollary 29 is therefore equivalent to Theorem 1 of Herings, Mauleon, and Vannetelbosch (2009) that characterizes the pairwise myopically stable set as the union of closed cycles. In their chapter, a closed cycle is defined in the sense of Jackson and Watts (2002) for network problems. The notion of closed cycle in Definition 13 is the appropriate generalization to social environments.

## 2.4.4 Normal-Form Games

As a final application, we consider normal-form games. We consider social environments where players are restricted to the use of pure strategies. The online supplement contains a formalization of the mixed extension with several results for that case.

A normal-form game  $G = (N, ((\Sigma_i, d_i), u_i)_{i \in N})$  consists of a set of players  $N$  and for each player  $i \in N$  a non-empty and compact metric space  $(\Sigma_i, d_i)$  of pure strategies and a utility function  $u_i : \Sigma \rightarrow \mathbb{R}$  over the set of strategy profiles  $\Sigma = \prod_{i \in N} \Sigma_i$ . A typical element of  $\Sigma$  is denoted by  $s$ .

For the corresponding social environment  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$ , we equate the state space  $X$  with the set of strategy profiles  $\Sigma$  and endow it with the product metric  $d(s, s') = \sum_{i \in N} d_i(s_i, s'_i)$ . The preferences  $(\geq_i)_{i \in N}$  are such that  $s \geq_i s'$  if and only if  $u_i(s) \geq u_i(s')$ .

We write  $(s_S, s_{-S})$  for the strategy profile where  $s_S$  is the list of strategies of players in coalition  $S \in \mathcal{N}$  and  $s_{-S}$  is the list of strategies of all other players, i.e.,  $s_S = (s_j)_{j \in S}$  and  $s_{-S} = (s_j)_{j \in N \setminus S}$ . With a slight abuse of notation, we write  $s_i$  and  $s_{-i} = (s_j)_{j \in N \setminus \{i\}}$  for single-player coalitions  $S = \{i\}$ .

It remains to specify the effectivity correspondence  $E$  of the social environment  $\Gamma$ . By allowing for all coalitional deviations, we describe coalitional normal-form games. For-

mally, for a coalition  $S \in \mathcal{N}$ , the effectivity correspondence is such that  $S \in E(s, s')$  if and only if  $s_{-S} = s'_{-S}$ . For this effectivity correspondence, a strategy profile  $s \in \Sigma$  is a strong Nash equilibrium (Aumann, 1959) if and only if it is in the core of the social environment  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$ . Thus, any strong Nash equilibrium must be in the MSS of the associated social environment.

Instead of the coalitional approach, we will henceforth restrict attention to the more frequently analyzed case of a non-cooperative game. In this case, the effectivity correspondence  $E$  only allows singletons to move. Formally, we have that  $S \in E(s, s')$  if and only if  $s_{-S} = s'_{-S}$  and  $|S| = 1$ . A non-cooperative normal-form game  $G = (N, ((\Sigma_i, d_i), u_i)_{i \in N})$  then induces a social environment  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$  that is identical to the social environment defined for normal-form games with coalitional moves, except that the effectivity correspondence  $E$  only allows for coalitions of size one.

A strategy profile  $s \in \Sigma$  is said to be a pure strategy Nash equilibrium of the game  $G$  if for every  $i \in N$  and for every  $s'_i \in \Sigma_i$  it holds that  $u_i(s) \geq u_i(s'_i, s_{-i})$ . Note that a strategy profile is a pure strategy Nash equilibrium if and only if it is in the core of the social environment  $\Gamma$ . Corollary 16 then shows that every pure strategy Nash equilibrium belongs to every MSS.

In the next step, we define pseudo-potential games (Dubey, Haimanko, and Zapechelnyuk, 2006) and show that the MSS coincides with the set of pure-strategy Nash equilibria for this class of games.

**Definition 30** (Pseudo-Potential Game). The game  $G = (N, ((\Sigma_i, d_i), u_i)_{i \in N})$  is a pseudo-potential game if there exists a continuous function  $P : \Sigma \rightarrow \mathbb{R}$  such that, for all  $i \in N$ , for all  $s \in \Sigma$ ,  $\arg \max_{s_i \in \Sigma_i} u_i(s_i, s_{-i}) \supseteq \arg \max_{s_i \in \Sigma_i} P(s_i, s_{-i})$ .

Pseudo-potential games generalize ordinal potential games (Monderer and Shapley, 1996) and best-response potential games (Voorneveld, 2000). Moreover, Dubey, Haimanko, and Zapechelnyuk (2006) show that the class of pseudo-potential games contains games of strategic complements or substitutes with aggregation such as Cournot oligopoly games. Jensen (2010) extends this result to generalized quasi-aggregative games.<sup>3</sup>

**Theorem 31.** Let  $G = (N, ((\Sigma_i, d_i), u_i)_{i \in N})$  be a pseudo-potential game and let  $\Gamma$  be the induced social environment. If the utility functions  $(u_i)_{i \in N}$  are continuous, then the MSS of  $\Gamma$  coincides with the set of pure strategy Nash equilibria of  $G$ .

For finite supermodular games, Friedman and Mezzetti (2001) establish the weak finite improvement property which implies the weak improvement property. Thus, the equivalence between the set of pure strategy Nash equilibria and the MSS also extends to this class of games.

## 2.5 Conclusion

The myopic stable set provides a solution concept for a wide variety of social environments. As we have shown, the setting encompasses coalition function form games, models

<sup>3</sup>Generalized quasi-aggregative games include aggregative games (Selten, 1970). For a subclass of aggregative games, the equivalence result in Theorem 31 can be obtained from Dindoš and Mezzetti (2006).



of network formation, matching models, and non-cooperative games. These environments have been chosen based on their prominence in the literature but are by no means exhaustive. In particular, promising environments for future research on the myopic stable set include exchange processes in general equilibrium models and many-to-many matching models with transfers.

The following three features boost the appeal of the myopic stable set as a solution concept. First, the myopic stable set unifies standard solution concepts in many social environments. For instance, it coincides with the coalition structure core in coalition function form games (Kóczy and Lauwers, 2004) if the coalition structure core is non-empty, the set of stable matchings in the one-to-one matching model of Gale and Shapley (1962), the set consisting of pairwise stable networks and closed cycles of networks (Jackson and Watts, 2002), and the set of pure strategy Nash equilibria in finite supermodular games (Topkis, 1979 and Milgrom and Roberts, 1990) and pseudo-potential games (Dubey, Haimanko, and Zapechelnyuk, 2006).

Second, our solution concept exists for any social environment and—under weak continuity assumptions—provides a unique set-valued prediction. This differs from well-known concepts in the literature which fail to satisfy these properties even in social environments with more structure.

Third, in important classes of problems, the MSS gives sharp predictions. For instance, for matching markets empirical findings starting with Roth and Peranson (1999) suggest that the core is small and Ashlagi, Kanoria, and Leshno (2017) provide theoretical arguments for why this is the case. The equivalence between the core of matching problems and the MSS then implies that MSS has significant predictive power. Other examples are the equivalence between MSS and the set of pure Nash equilibria pseudo-potential games and finite supermodular games.

At the same time, there are cases where the MSS may be large. For example, when the MSS contains states that do not belong to the core as in the matching model by Knuth (1976) as discussed in Subsection 2.4.2. Intuitively, the combination of myopic behavior and a restrictive effectivity correspondence may result in cycling and hence, a large MSS. This reflects the trade-off between a general solution concept for which existence and non-emptiness is guaranteed, like the MSS, and a clear prediction for every class of social environments. The investigation of refinements of the MSS in such cases is a natural direction for future research.

## 2.6 Appendix

**PROOF OF THEOREM 6:** First observe that the set of states  $X$  is a QMSS. Indeed, since it is compact, it is closed and it trivially satisfies deterrence of external deviations and asymptotic external stability.

Let  $\mathcal{Z}$  be the collection of all sets of states that are a QMSS. Notice that  $\mathcal{Z}$  is non-empty as  $X \in \mathcal{Z}$ . A set  $Z' \in \mathcal{Z}$  is a maximal element in the partially ordered set  $(\mathcal{Z}, \supseteq)$  if for all  $Z \in \mathcal{Z}$  with  $Z' \supseteq Z$ , we have  $Z = Z'$ . We will use Zorn's lemma to show the existence of a maximal element in the partially ordered set  $(\mathcal{Z}, \supseteq)$ .

Let  $\mathcal{S}$  be a chain in  $\mathcal{Z}$ , i.e.,  $(\mathcal{S}, \supseteq)$  is a totally ordered subset of  $(\mathcal{Z}, \supseteq)$ . Let  $I$  be an index set for the sets in  $\mathcal{S}$ , i.e.,  $\mathcal{S} = \{Z^\alpha | \alpha \in I\}$ . Let  $\triangleright$  be the order on  $I$  that is induced by the

order on  $\mathcal{S}$ , i.e.,  $\beta \triangleright \alpha$  if and only if  $Z^\alpha \supseteq Z^\beta$ . In order to apply Zorn's Lemma, we have to show that  $\mathcal{S}$  has an upper bound in  $\mathcal{Z}$ . Let  $M = \bigcap_{\alpha \in I} Z^\alpha$ . Clearly,  $M$  is an upper bound of  $\mathcal{S}$ . We proceed by showing that  $M \in \mathcal{Z}$ , i.e.,  $M$  is a QMSS. First of all, observe that  $M$  is closed as it is defined as the intersection of a collection of closed sets. We need to show that it satisfies deterrence of external deviations and asymptotic external stability.

**Deterrence of external deviations:** Let  $x \in M$  and  $y \notin M$  be given. Then there is  $\alpha \in I$  such that  $y \notin Z^\alpha$ , since otherwise  $y \in Z^\alpha$  for all  $\alpha \in I$ , which means that  $y \in M$ . Since  $x \in Z^\alpha$  and  $Z^\alpha$  satisfies deterrence of external deviations, we obtain  $y \notin f(x)$  as was to be shown.

**Asymptotic external stability:** Consider some  $y \notin M$ . Then there is  $\alpha \in I$  such that  $y \notin Z^\alpha$ . As  $\mathcal{S}$  is a chain, it follows that for all  $\beta \triangleright \alpha$  we have  $y \notin Z^\beta$ .

For every  $\beta \triangleright \alpha$ , there is  $x^\beta \in Z^\beta$  such that  $x^\beta \in f^\infty(y)$ , since  $Z^\beta$  satisfies asymptotic external stability. This defines a net  $(x^\beta)_{\beta \triangleright \alpha}$ . Given that  $X$  is compact, it follows by Theorem 2.31 of Aliprantis and Border (2006) that this net has a convergent subnet, say  $(x^{\beta'})_{\beta' \in I'}$ , where  $I' \subseteq I$  is such that for all  $\beta \in I$  there is a  $\beta' \in I'$  such that  $\beta' \triangleright \beta$ . Let  $\bar{x}$  be the limit of this convergent subnet. We split the remaining part of the proof in two steps. First, we show that  $\bar{x} \in M$ . Second, we show that  $\bar{x} \in f^\infty(y)$ .

**Step 1:  $\bar{x} \in M$ :** Towards a contradiction, suppose that  $\bar{x} \notin M$ . Then, there exists  $\gamma \in I$  such that  $\bar{x} \notin Z^\gamma$ . In particular, given that  $Z^\gamma$  is a closed set, there is  $\varepsilon > 0$  such that  $B_\varepsilon(\bar{x}) \cap Z^\gamma = \emptyset$ . Since  $\mathcal{S}$  is a chain, we have that  $B_\varepsilon(\bar{x}) \cap Z^\delta = \emptyset$  for all  $\delta \triangleright \gamma$ . Since  $\bar{x}$  is the limit of the subnet  $(x^{\beta'})_{\beta' \in I'}$ , there is  $\gamma' \in I'$  such that  $\gamma' \triangleright \gamma$  and  $x^{\gamma'} \in B_\varepsilon(\bar{x})$ . Then we have  $x^{\gamma'} \in Z^{\gamma'}$ ,  $x^{\gamma'} \in B_\varepsilon(\bar{x})$ , and  $B_\varepsilon(\bar{x}) \cap Z^{\gamma'} = \emptyset$ , a contradiction. We conclude that  $\bar{x} \in M$ .

**Step 2:  $\bar{x} \in f^\infty(y)$ :** We need to show that for every  $\varepsilon > 0$  there is  $k \in \mathbb{N}$  and  $x \in f^k(y)$  such that  $x \in B_\varepsilon(\bar{x})$ .

Let some  $\varepsilon > 0$  be given. The subnet  $(x^{\beta'})_{\beta' \in I'}$  converges to  $\bar{x}$ . As such, there exists  $\gamma' \in I'$  such that  $x^{\gamma'} \in B_{\varepsilon/2}(\bar{x})$ . In addition,  $x^{\gamma'} \in f^\infty(y)$ , so there is  $k \in \mathbb{N}$  and  $x \in f^k(y)$  such that  $x \in B_{\varepsilon/2}(x^{\gamma'})$ . Then, by the triangle inequality, it holds that  $x \in B_\varepsilon(\bar{x})$ . Together with  $x \in f^k(y)$ , this concludes the proof, i.e.,  $\bar{x} \in f^\infty(y)$ .  $\square$

**PROOF OF LEMMA 7:** Step 1: Let  $\Gamma$  be a social environment and let  $M$  be an MSS of  $\Gamma$ . We show in Step 1 that if  $x \in M$  and  $y \in f^\infty(x)$ , then  $y \in M$ .

Let  $x \in M$  and  $y \in f^\infty(x)$  and suppose, towards a contradiction, that  $y \notin M$ . Given that  $M$  is closed, there is  $\varepsilon > 0$  such that  $B_\varepsilon(y) \cap M = \emptyset$ . Also, by definition, there is  $k \in \mathbb{N}$  and  $z \in f^k(x)$  such that  $z \in B_\varepsilon(y)$ , i.e.,  $z \notin M$ . Since  $z \in f^k(x)$ , there is a sequence  $z^0, z^1, \dots, z^k$  of length  $k$  such that

$$z^0 = x, z^1 \in f(z^0), \dots, z^k = z \in f(z^{k-1}).$$

Let  $k' \in \{1, \dots, k\}$  be such that  $z^{k'}$  is the first element in this sequence with the property that  $z^{k'} \notin M$ . Given that  $z^0 = x \in M$  and  $z^k = z \notin M$ , such an element exists. It holds that  $z^{k'-1} \in M$ ,  $z^{k'} \in f(z^{k'-1})$ , and  $z^{k'} \notin M$ . This contradicts deterrence of external deviations for  $M$ .

Step 2: It remains to show that  $M_1 \cap M_2 \neq \emptyset$ .

Consider a state  $x_1 \in M_1$ . If  $x_1 \in M_2$ , then we are done. Otherwise, by asymptotic external stability of  $M_2$ , we know that there is  $x_2 \in M_2$  such that  $x_2 \in f^\infty(x_1)$ . By Step 1,

we have that  $x_2 \in M_1$ , so  $x_2 \in M_1 \cap M_2$ .  $\square$

The following technical lemma is helpful in proving Theorem 9.

**Lemma 32.** If the dominance correspondence  $f : X \rightarrow X$  is lower hemi-continuous, then the asymptotic dominance correspondence  $f^\infty : X \rightarrow X$  is transitive.

*Proof.* Let  $x, y, z \in X$  be such that  $y \in f^\infty(x)$  and  $z \in f^\infty(y)$ . We have to show that  $z \in f^\infty(x)$ , so we need to show that for every  $\varepsilon > 0$ , there is  $k' \in \mathbb{N}$  and  $z' \in f^{k'}(x)$  such that  $z' \in B_\varepsilon(z)$ .

By assumption,  $z \in f^\infty(y)$ , so there is  $k \in \mathbb{N}$  and  $z_1 \in f^k(y)$  such that  $z_1 \in B_{\varepsilon/2}(z)$ . In addition, as  $y \in f^\infty(x)$ , we know that for every  $\ell \in \mathbb{N}$  there is  $k_\ell \in \mathbb{N}$  and  $y^\ell \in f^{k_\ell}(x)$  such that  $y^\ell \in B_{1/\ell}(y)$ . This generates a sequence  $(y^\ell)_{\ell \in \mathbb{N}}$  that converges to  $y$ , i.e.,  $y^\ell \rightarrow y$ .

Note that  $f^k$  is lower hemi-continuous, since it is a composition of  $k$  lower hemi-continuous correspondences. Given lower hemi-continuity of  $f^k$  and the fact that  $z_1 \in f^k(y)$ , we know that there is a sequence  $(z_2^\ell)_{\ell \in \mathbb{N}}$  such that  $z_2^\ell \rightarrow z_1$  and  $z_2^\ell \in f^k(y^\ell)$ . Now, we have that  $y^\ell \in f^{k_\ell}(x)$  and  $z_2^\ell \in f^k(y^\ell)$ , which gives  $z_2^\ell \in f^{k+k_\ell}(x)$ .

We take  $\ell$  large enough such that  $z_2^\ell \in B_{\varepsilon/2}(z_1)$ . Since  $z_1 \in B_{\varepsilon/2}(z)$ , the triangular inequality gives  $z_2^\ell \in B_\varepsilon(z)$ . This completes the proof.  $\square$

**PROOF OF THEOREM 9:** Suppose not, then, by Theorem 6 and Lemma 7, there exists an MSS  $M_1$  and an MSS  $M_2$  such that  $M_1 \neq M_2$  and their intersection  $M_3 = M_1 \cap M_2$  is non-empty. Let us show that  $M_3$  is a QMSS, contradicting the minimality of  $M_1$  and  $M_2$ , and establishing the uniqueness of the MSS. First of all, notice that  $M_3$ , being the intersection of two closed sets, is also closed.

For deterrence of external deviations, let  $x \in M_3$  and  $y \in f(x)$ . Then given that  $x \in M_1$  and  $M_1$  satisfies deterrence of external deviations, it must be that  $y \in M_1$ . Also given that  $x \in M_2$  and  $M_2$  satisfies deterrence of external deviations, it must be that  $y \in M_2$ . This implies that  $y \in M_1 \cap M_2 = M_3$  as was to be shown.

For asymptotic external stability, take any  $y \notin M_3$ . There are three cases to consider.

**Case 1:  $y \in M_1 \setminus M_3$ :** Then, by asymptotic external stability of  $M_2$ , there is  $x \in M_2$  such that  $x \in f^\infty(y)$ . By Step 1 in the proof of Lemma 7, we have that  $x \in M_1$ . This means that  $x \in M_1 \cap M_2 = M_3$ , which is what we needed to show.

**Case 2:  $y \in M_2 \setminus M_3$ :** The proof is symmetric to Case 1 with  $M_1$  and  $M_2$  interchanged.

**Case 3:  $y \in X \setminus (M_1 \cup M_2)$ :** We know, by asymptotic external stability of  $M_1$ , that there is  $x \in M_1$  such that  $x \in f^\infty(y)$ . If  $x \in M_3$ , we are done. If not, we know from Case 1 above that there is  $z \in M_3$  such that  $z \in f^\infty(x)$ . It follows from  $x \in f^\infty(y)$  and  $z \in f^\infty(x)$  that  $z \in f^\infty(y)$  by Lemma 32.  $\square$

**PROOF OF THEOREM 12:** Let  $x, y \in X$  and sequences  $(x^k)_{k \in \mathbb{N}}$  and  $(y^k)_{k \in \mathbb{N}}$  in  $X$  such that  $x^k \rightarrow x$  and  $y^k \rightarrow y$  be given. Let us first show that if individual  $i \in N$  strictly prefers  $y$  to  $x$ , so  $y \succ_i x$ , then there is  $\ell \in \mathbb{N}$  such that for all  $k \geq \ell$ ,  $y^k \succ_i x^k$ . Suppose not, then for every  $\ell \in \mathbb{N}$  we can find  $k_\ell \geq \ell$  such that  $x^{k_\ell} \succeq_i y^{k_\ell}$ . This creates sequences  $(x^{k_\ell})_{\ell \in \mathbb{N}}$ ,  $(y^{k_\ell})_{\ell \in \mathbb{N}}$  in  $X$  with  $x^{k_\ell} \rightarrow x$  and  $y^{k_\ell} \rightarrow y$  such that, for every  $\ell \in \mathbb{N}$ ,  $x^{k_\ell} \succeq_i y^{k_\ell}$ . By continuity of  $\succeq_i$ , it holds that  $x \succeq_i y$ , a contradiction.

Let  $(x^k)_{k \in \mathbb{N}}$  be a sequence in  $X$  such that  $x^k \rightarrow x$  and consider some  $y \in f(x)$ . Then either  $y = x$  or  $y \neq x$  and there is a coalition  $S \in E(x, y)$  such that, for every  $i \in S$ ,  $y \succ_i x$ .

If  $y = x$ , take the sequence  $(y^k)_{k \in \mathbb{N}}$  in  $X$  defined by  $y^k = x^k$ . We immediately have that, for every  $k \in \mathbb{N}$ ,  $y^k \in f(x^k)$ , and  $y^k \rightarrow y$ .

We now consider the case where  $y \neq x$  and there is a coalition  $S \in E(x, y)$  such that, for every  $i \in S$ ,  $y >_i x$ . By lower hemi-continuity of the correspondence  $G_S$ , we know that there is a sequence  $(y^k)_{k \in \mathbb{N}}$  such that  $y^k \in G_S(x^k)$  and  $y^k \rightarrow y$ . By the first paragraph of the proof, we know that for every  $i \in S$  there is  $\ell_i \in \mathbb{N}$  such that, for every  $k \geq \ell_i$ ,  $y^k >_i x^k$ . Let  $\ell = \max_{i \in S} \ell_i$ . Then, for every  $k \geq \ell$ , for every  $i \in S$ ,  $y^k >_i x^k$ , and  $S \in E(x^k, y^k)$ , which shows that  $y^k \in f(x^k)$ . The sequence  $(z^k)_{k \in \mathbb{N}}$  defined by  $z^k = x^k$  if  $k < \ell$  and  $z^k = y^k$  if  $k \geq \ell$  therefore has all the desired properties: for every  $k \in \mathbb{N}$ ,  $z^k \in f(x^k)$ , and  $z^k \rightarrow y$ .

As  $f$  is lower hemi-continuous, Theorem 9 implies that the social environment has a unique MSS.  $\square$

**PROOF OF THEOREM 14:** Towards a contradiction, suppose there is a closed cycle  $C$  which is not a subset of  $M$ . Let  $x \in C$  and  $x \notin M$ . By asymptotic external stability there is  $y \in M$  such that  $y \in f^\infty(x)$ . By definition of a closed cycle, it follows that  $y \in C$ . As  $x \in C$ , again by definition of a closed cycle, we also have that  $x \in f^\infty(y)$ . By Step 1 in the proof of Lemma 7, it follows that  $x \in M$ , a contradiction. Since the choice of  $C$  was arbitrary, we have shown that  $\text{CC} \subseteq M$ .

We show next that if  $X$  is finite, then  $\text{CC} = M$ . Since  $\text{CC} \subseteq M$ , we only need to show that  $\text{CC}$  is a QMSS. The set  $\text{CC}$  satisfies deterrence of external deviations, since for all  $x \in \text{CC}$ ,  $f(x) \subseteq f^\infty(x) \subseteq \text{CC}$ . It remains to verify asymptotic external stability of  $\text{CC}$ , i.e., for every state  $x \notin \text{CC}$ ,  $f^\infty(x) \cap \text{CC} \neq \emptyset$ .

Let  $x \notin \text{CC}$  and define  $Y = f^\infty(x)$ . Note that  $Y$  is non-empty since  $x \in f(x)$ . It also holds that  $Y$  is finite and  $f^\infty(y) \subseteq Y$  for every  $y \in Y$ . Let us represent the set  $Y$  and the dominance correspondence  $f$  on  $Y$  by a finite directed graph  $D$ , i.e., (i)  $Y$  is the set of vertices of  $D$  and (ii)  $D$  has an arc from  $y$  to  $z$  if and only if  $z \in f(y)$ . By contracting each strongly connected component of  $D$  to a single vertex, we obtain a directed acyclic graph, which is called the condensation of  $D$ . As the condensation is finite and acyclic, it has a maximal element, say  $c$ . Observe that  $c$  represents a closed cycle  $C$ , so  $Y \cap \text{CC} \neq \emptyset$ .  $\square$

**PROOF OF THEOREM 18:** Assume that  $\Gamma$  satisfies the weak improvement property. Let  $M$  be an MSS of  $\Gamma$ . By Corollary 16, we have  $\text{CO} \subseteq M$ . We will show that  $\text{CO}$  is a QMSS. By minimality, it then follows that  $\text{CO} = M$ .

In order to see that  $\text{CO}$  is closed, let  $(x^k)_{k \in \mathbb{N}}$  be a sequence in  $\text{CO}$ , so for all  $k \in \mathbb{N}$  it holds that  $f(x^k) = \{x^k\}$ . Now assume that  $x^k \rightarrow x$  and  $x \notin \text{CO}$ . This means that there is  $y \neq x$  such that  $y \in f(x)$ . By lower hemi-continuity of  $f$ , there is a sequence  $(y^k)_{k \in \mathbb{N}}$  such that  $y^k \in f(x^k)$  and  $y^k \rightarrow y$ . As for all  $k \in \mathbb{N}$ ,  $x^k \in \text{CO}$ , we have that  $y^k = x^k$ , which means that  $y^k \rightarrow x$ , so  $y = x$ , a contradiction. Deterrence of external deviations is immediate for the core as it is the union of the sinks. If the social environment satisfies the weak improvement property, we have that for all  $x \notin \text{CO}$ ,  $f^\infty(x) \cap \text{CO} \neq \emptyset$ , thus the core satisfies asymptotic external stability.

For the reverse, let  $M$  be an MSS equal to  $\text{CO}$ . Now, if  $x \in M$ , it is a sink, so  $f^\infty(x) = \{x\} \subseteq \text{CO}$ . If  $x \notin \text{CO}$  we have by asymptotic external stability of  $M$  that  $f^\infty(x) \cap M \neq \emptyset$ , so  $f^\infty(x)$  contains a sink, i.e.,  $\Gamma$  satisfies the weak improvement property.  $\square$

**PROOF OF THEOREM 20:** Suppose that  $M$  is an MSS such that  $M \cap V = \emptyset$ . Recall that  $M$

is non-empty. Let  $x \in M$  and  $x \notin V$ . Since  $V$  satisfies external stability, there is  $y \in f(x)$  such that  $y \in V$ . Since  $M$  satisfies deterrence of external deviations, we have that  $y \in M$ , so  $M \cap V \neq \emptyset$ , a contradiction.  $\square$

**PROOF OF THEOREM 21:** By Theorem 14, the MSS is unique and equal to the union of all closed cycles  $CC$ . Since  $Q$  is consistent with  $f$ , a state is recurrent if and only if it belongs to a closed cycle.  $\square$

**PROOF OF THEOREM 25:**

**Step 1:  $\Gamma$  has a unique MSS.**

We show lower hemi-continuity of  $E$  and continuity of the preferences. The result then follows from Theorem 12.

To show lower hemi-continuity of  $E$ , let some  $S \in \mathcal{N}$ , a sequence  $(x^k)_{k \in \mathbb{N}}$  in  $X$  such that  $x^k \rightarrow x$ , and some  $y \in G_S(x)$  be given. We show that there is a sequence  $(y^k)_{k \in \mathbb{N}}$  such that  $y^k \in G_S(x^k)$  and  $y^k \rightarrow y$ . If  $y = x$ , then the choice  $y^k = x^k$  would do, so consider the case  $y \neq x$ .

First of all, there is  $k' \in \mathbb{N}$  such that for all  $k \geq k'$ ,  $\pi(x^k) = \pi(x)$ , so in particular  $\mathcal{U}(x^k, S) = \mathcal{U}(x, S)$ . For every  $k < k'$ , we define  $y^k = x^k$ . For every  $k \geq k'$ , we define  $y^k \in X$  by  $\pi(y^k) = \pi(y)$  and

$$u_i(y^k) = \begin{cases} u_i(y), & i \in N \setminus U(x, S), \\ u_i(x^k), & i \in U(x, S). \end{cases}$$

This completely specifies the state  $y^k$ . Consider some  $k \geq k'$ . Since  $y \neq x$  and  $y \in G_S(x)$ , it holds that  $S \in \pi(y)$  and, for every  $i \in N \setminus (S \cup U(x, S))$ , we have that  $i$  is a residual player. The properties of the  $\gamma$ -model imply that  $\{i\} \in \pi(y)$ . The same properties hold for  $\pi(y^k)$ . For every  $i \in S$ , it holds that  $u_i(y^k) = u_i(y)$ , so  $u_i(y^k) \geq v(\{i\})$  and  $\sum_{i \in S} u_i(y^k) = v(S)$ . For every  $i \in N \setminus (S \cup U(x, S))$ , we have that  $u_i(y^k) = v(\{i\}) = u_i(y)$ . For every  $i \in U(x, S)$  it holds that  $u_i(y) = u_i(x)$  and  $u_i(y^k) = u_i(x^k)$ . By coalitional sovereignty, we have that  $y^k \in G_S(x^k)$ . Using that  $x^k \rightarrow x$ , it follows easily that  $y^k \rightarrow y$ .

Let some  $i \in N$  be given. To show continuity of  $\geq_i$ , let  $(x^k)_{k \in \mathbb{N}}$  and  $(y^k)_{k \in \mathbb{N}}$  be sequences in  $X$  such that  $x^k \rightarrow x$  and  $y^k \rightarrow y$ . Then, by continuity of the projection, we have that  $u_i(x^k) \rightarrow u_i(x)$  and  $u_i(y^k) \rightarrow u_i(y)$ . So if  $u_i(x^k) \geq u_i(y^k)$  for all  $k \in \mathbb{N}$ , we obtain  $u_i(x) \geq u_i(y)$ , which shows that  $x \geq_i y$ .

**Step 2: The core of  $\Gamma$  is equal to  $Y$ .**

Let  $y \in CO$  and suppose  $y \notin Y$ . Then there is a coalition  $S \in \mathcal{N}$  such that  $\sum_{i \in S} u_i(y) < v(S)$ . Since  $y \in X$ , it holds for all  $i \in S$  that  $u_i(y) \geq v(\{i\})$ . Now, let  $u_S$  be a vector of payoffs for the members in  $S$  such that  $\sum_{i \in S} u_i = v(S)$  and for all  $i \in S$ ,  $u_i > u_i(y)$ . Then, by full support, there exists a state  $y' \in X$  such that  $S \in E(y, y')$  and  $u_S = u_S(y')$ . We conclude that  $y' \in f(y)$ . This contradicts the fact that  $y \in CO$ .

For the reverse, let  $y \in Y$  and suppose there is  $z \in f(y)$  such that  $z \neq y$ , i.e.,  $y \notin CO$ . Then there is  $S \in E(y, z)$  such that for all  $i \in S$  it holds that  $u_i(z) > u_i(y)$ . Also,  $v(S) = \sum_{i \in S} u_i(z) > \sum_{i \in S} u_i(y) \geq v(S)$ , where the first equality follows from the definition of the state space and the last inequality from the definition of  $Y$ . We have obtained a contradiction.

**Step 3: If  $Y \neq \emptyset$ , then the MSS of  $\Gamma$  is equal to  $Y$ .**

Step 2 shows that  $Y$  is equal to the core of  $\Gamma$ . If we can show that  $\Gamma$  satisfies the weak

improvement property whenever  $Y \neq \emptyset$ , then we can use Theorem 18 to establish our result. Since the proof is trivial when the number of individuals  $n$  is equal to 1, we assume  $n \geq 2$  throughout.

Assume that  $Y \neq \emptyset$ . We need to show that for all  $x^0 \in X$ ,  $f^\infty(x^0) \cap Y \neq \emptyset$ . If  $x^0$  in  $Y$ , then nothing needs to be shown, so assume that  $x^0 \in X \setminus Y$ . We need to show that for every  $\varepsilon > 0$  there is a number  $k' \in \mathbb{N}$ , a state  $x^{k'} \in f^{k'}(x^0)$ , and a state  $y \in Y$  such that  $d(x^{k'}, y) < \varepsilon$ .

Let some  $\varepsilon > 0$  be given. Béal, Rémila, and Solal (2013) show that there exists a sequence of states  $(x^0, \dots, x^{k'})$  such that  $x^{k'} \in Y$ ,  $k'$  is less than or equal to  $(n^2 + 4n)/4$ , and, for every  $k \in \{1, \dots, k'\}$ ,

1. there is  $S^k \in \mathcal{N}$  such that  $S^k \in E(x^{k-1}, x^k)$ ,
2.  $u_{S^k}(x^{k-1}) < u_{S^k}(x^k)$ .

Notice that the inequality in 2. only means that at least one of the players in  $S^k$  gets a strictly higher payoff, not necessarily all of them. Let  $P^k$  be the set of partners of the players in  $S^k$  at state  $x^{k-1}$ , more formally defined as

$$P^k = \cup_{\{S \in \pi(x^{k-1}) \mid S \cap S^k \neq \emptyset\}} S,$$

so  $P^k$  is equal to the moving coalition  $S^k$  together with the residual players. Since  $S^k \in E(x^{k-1}, x^k)$ , it follows that

$$\begin{aligned} u_i(x^k) &= v(\{i\}), & i \in P^k \setminus S^k, \\ u_i(x^k) &= u_i(x^{k-1}), & i \in N \setminus P^k. \end{aligned}$$

We define  $W^k \subset S^k$  to be the, possibly empty, proper subset of  $S^k$  consisting of players that only weakly improve when moving from state  $x^{k-1}$  to state  $x^k$ , so for every  $i \in W^k$  it holds that  $u_i(x^{k-1}) = u_i(x^k)$ . We define

$$\delta = \min_{k \in \{1, \dots, k'\}} \min_{i \in S^k \setminus W^k} u_i(x^k) - u_i(x^{k-1}) \text{ and } \varepsilon' = \min\{\delta, \varepsilon\},$$

so  $\delta$  is the smallest improvement of any of the strictly improving players involved in any move along the sequence. It holds that  $\delta > 0$  and therefore that  $\varepsilon' > 0$ . For  $k \in \{0, \dots, k'\}$ , define  $\nu_k = n^{2k}/(n^{2k'+1})$ . We define  $e(W^k) = 0$  if  $W^k = \emptyset$  and  $e(W^k) = 1$  otherwise. We use the sequence  $(x^0, x^1, \dots, x^{k'})$  of states as constructed by Béal, Rémila, and Solal (2013) to define a new sequence  $(\tilde{x}^0, \tilde{x}^1, \dots, \tilde{x}^{k'})$  of states by setting  $\tilde{x}^0 = x^0$  and, for every  $k \in \{1, \dots, k'\}$ ,

$$\begin{aligned} \pi(\tilde{x}^k) &= \pi(x^k), \\ u_i(\tilde{x}^k) &= u_i(x^k) + \varepsilon' \nu_k \frac{|S^k \setminus W^k|}{|W^k|}, & i \in W^k, \\ u_i(\tilde{x}^k) &= u_i(x^k) - \varepsilon' \nu_k e(W^k), & i \in S^k \setminus W^k, \\ u_i(\tilde{x}^k) &= u_i(x^k) = v(\{i\}), & i \in P^k \setminus S^k, \\ u_i(\tilde{x}^k) &= u_i(\tilde{x}^{k-1}), & i \in N \setminus P^k. \end{aligned}$$

Notice that the first line does not entail a division by zero, since if  $i \in W^k$ , then  $W^k \neq \emptyset$ .

Compared to the sequence  $(x^0, x^1, \dots, x^{k'})$ , the sequence  $(\tilde{x}^0, \tilde{x}^1, \dots, \tilde{x}^{k'})$  is such that each strictly improving player in  $S^k \setminus W^k$  donates an amount  $\varepsilon' \nu_k / |W^k|$  to each of the players

in  $W^k$  whenever the latter set is non-empty. It is also important to observe that the fraction  $\nu_k$  is an  $n^2$  multiple of  $\nu_{k-1}$  and that  $\nu_{k'} = 1/n$ .

We show first by induction that, for every  $k \in \{0, \dots, k'\}$ ,  $\tilde{x}^k \in X$ . Obviously, it holds that  $\tilde{x}^0 = x^0 \in X$ . Assume that, for some  $k \in \{1, \dots, k'\}$ ,  $\tilde{x}^{k-1} \in X$ . We show that  $\tilde{x}^k \in X$ . It holds that

$$\begin{aligned} u_i(\tilde{x}^k) &> u_i(x^k) \geq v(\{i\}), & i \in W^k, \\ u_i(\tilde{x}^k) &\geq u_i(x^{k-1}) + \delta - \varepsilon' \nu_k > u_i(x^{k-1}) + \delta - \varepsilon' \geq u_i(x^{k-1}) \geq v(\{i\}), & i \in S^k \setminus W^k, \\ u_i(\tilde{x}^k) &= v(\{i\}), & i \in P^k \setminus S^k, \\ u_i(\tilde{x}^k) &= u_i(\tilde{x}^{k-1}) \geq v(\{i\}), & i \in N \setminus P^k, \end{aligned}$$

where the very last inequality follows from the induction hypothesis. Moreover, for every  $S \in \pi(x^k)$ , it holds that either  $S = S^k$  and  $W^k = \emptyset$ , so

$$\sum_{i \in S} u_i(\tilde{x}^k) = \sum_{i \in S^k} u_i(x^k) = v(S),$$

or  $S = S^k$  and  $W^k \neq \emptyset$ , so

$$\sum_{i \in S} u_i(\tilde{x}^k) = \sum_{i \in W^k} \left( u_i(x^k) + \varepsilon' \nu_k \frac{|S^k \setminus W^k|}{|W^k|} \right) + \sum_{i \in S^k \setminus W^k} (u_i(x^k) - \varepsilon' \nu_k) = \sum_{i \in S^k} u_i(x^k) = v(S),$$

or  $S = \{i'\}$  with  $i' \in P^k \setminus S^k$  and

$$\sum_{i \in S} u_i(\tilde{x}^k) = u_{i'}(\tilde{x}^k) = u_{i'}(x^k) = v(\{i'\}) = v(S),$$

or  $S \subseteq N \setminus P^k$ , so  $S \in \pi(\tilde{x}^{k-1})$ , and

$$\sum_{i \in S} u_i(\tilde{x}^k) = \sum_{i \in S} u_i(\tilde{x}^{k-1}) = v(S),$$

where the last equality makes use of the induction hypothesis. We have now completed the proof of the fact that for every  $k \in \{0, \dots, k'\}$ ,  $\tilde{x}^k \in X$ .

We show next by induction that, for every  $k \in \{0, \dots, k'\}$ , and for every  $i \in N$ ,

$$|u_i(\tilde{x}^k) - u_i(x^k)| \leq \varepsilon' \nu_k (n - 1).$$

Obviously, for every  $i \in N$ , it holds that  $|u_i(\tilde{x}^0) - u_i(x^0)| = 0 \leq \varepsilon' \nu_0 (n - 1)$ . Assume that, for some  $k \in \{1, \dots, k'\}$ , for every  $i \in N$ ,  $|u_i(\tilde{x}^{k-1}) - u_i(x^{k-1})| \leq \varepsilon' \nu_{k-1} (n - 1)$ . We show that, for every  $i \in N$ ,  $|u_i(\tilde{x}^k) - u_i(x^k)| \leq \varepsilon' \nu_k (n - 1)$ . If  $i \in W^k$ , then  $W^k \neq \emptyset$ , and the statement follows from the observation that

$$0 \leq u_i(\tilde{x}^k) - u_i(x^k) = \varepsilon' \nu_k \frac{|S^k \setminus W^k|}{|W^k|} \leq \varepsilon' \nu_k (n - 1).$$

If  $i \in S^k \setminus W^k$ , then we have that

$$0 \geq u_i(\tilde{x}^k) - u_i(x^k) \geq -\varepsilon' \nu_k \geq -\varepsilon' \nu_k (n - 1).$$

If  $i \in P^k \setminus S^k$ , then we have  $|u_i(\tilde{x}^k) - u_i(x^k)| = 0$ . If  $i \in N \setminus P^k$ , then it holds that

$$|u_i(\tilde{x}^k) - u_i(x^k)| = |u_i(\tilde{x}^{k-1}) - u_i(x^{k-1})| \leq \varepsilon' \nu_{k-1}(n-1) < \varepsilon' \nu_k(n-1),$$

where the first inequality makes use of the induction hypothesis and the last inequality of the fact that  $\nu_{k-1} < \nu_k$ .

Let some  $k \in \{1, \dots, k'\}$  and some  $i \in S^k$  be given. We show that  $u_i(\tilde{x}^k) > u_i(\tilde{x}^{k-1})$ . If  $i \in W^k$ , then it holds that

$$\begin{aligned} u_i(\tilde{x}^k) &= u_i(x^k) + \varepsilon' \nu_k \frac{|S^k \setminus W^k|}{|W^k|} = u_i(x^{k-1}) + \varepsilon' \nu_k \frac{|S^k \setminus W^k|}{|W^k|} \\ &\geq u_i(\tilde{x}^{k-1}) - \varepsilon' \nu_{k-1}(n-1) + \varepsilon' \nu_k \frac{1}{n-1} > u_i(\tilde{x}^{k-1}), \end{aligned}$$

where the strict inequality uses that  $\nu_k = n^2 \nu_{k-1}$ . If  $i \in S^k \setminus W^k$ , then it holds that

$$\begin{aligned} u_i(\tilde{x}^k) &\geq u_i(x^k) - \varepsilon' \nu_k \geq u_i(x^{k-1}) + \delta - \varepsilon' \nu_k \\ &\geq u_i(\tilde{x}^{k-1}) - \varepsilon' \nu_{k-1}(n-1) + \delta - \varepsilon' n^2 \nu_{k-1} > u_i(\tilde{x}^{k-1}), \end{aligned}$$

where the strict inequality uses the facts that  $\delta \geq \varepsilon'$  and

$$(n^2 + (n-1))\nu_{k-1} < 2n^2 \nu_{k-1} \leq 2\nu_k \leq 1.$$

Combining the statements proven so far, it follows that  $\tilde{x}^{k'} \in f^{k'}(x^0)$ . We complete the proof of the weak improvement property by noting that  $x^{k'} \in Y$  by the result of Béal, Rémila, and Solal (2013) and by observing that  $d(\tilde{x}^{k'}, x^{k'}) < \varepsilon$  since  $\pi(\tilde{x}^{k'}) = \pi(x^{k'})$  and, for every  $i \in N$ ,

$$|u_i(\tilde{x}^{k'}) - u_i(x^{k'})| \leq \varepsilon' \nu_{k'}(n-1) < \varepsilon' \leq \varepsilon.$$

□

**PROOF OF THEOREM 26:** We show first that  $F(X)$  satisfies deterrence of external deviations. Let  $x \in F(X)$  and  $y \in f(x)$  be given and let  $S \in E(x, y)$  be such that, for every  $i \in S$ ,  $u_i(y) > u_i(x)$ . It follows that  $S \in \mathcal{W}$ . Since  $\sum_{i \in N} u_i(x) = \sum_{i \in N} u_i(y) = 1$ , we have that  $S \neq N$ . Since  $N \setminus S \notin \mathcal{W}$ ,  $S$  has a non-empty intersection with the unique coalition in  $\pi(x) \cap \mathcal{W}$ . It follows that  $\pi(y) = \{S\} \cup \{\{i\} \mid i \in N \setminus S\}$ , so it holds that  $y \in F(X)$ .

We show next that  $F(X)$  satisfies asymptotic external stability. Let some  $x \in X \setminus F(X)$  be given. If  $\pi(x) \cap \mathcal{W} \neq \emptyset$ , then let  $j \in N$  be such that  $u_j(x) > 0$ . If  $\pi(x) \cap \mathcal{W} = \emptyset$ , then take any  $j \in N$ . Since the core of  $v$  is empty, it holds that  $N \setminus \{j\} \in \mathcal{W}$ , as otherwise  $j$  would be a veto player. Let  $y \in F(X)$  be such that  $\pi(y) = \{N \setminus \{j\}, \{j\}\}$  and  $u_{N \setminus \{j\}}(y) \gg u_{N \setminus \{j\}}(x)$ . Since  $N \setminus \{j\} \in E(x, y)$  it holds that  $y \in f(x)$ .

We show next that for every  $x \in F(X)$  it holds that  $f^\infty(x) = F(X)$ . Let some  $x, y \in F(X)$  be given and denote the winning coalition in  $\pi(y)$  by  $W$ . We have to show that for every  $\varepsilon > 0$  there exists  $k' \in \mathbb{N}$  and  $z \in f^{k'}(x)$  such that  $z \in B_\varepsilon(y)$ .

Let some  $\varepsilon \in (0, 1/n)$  and  $S \in \pi(y) \setminus \{W\}$  be given. We show first that there is  $k \in \mathbb{N}$  and  $x^k \in f^k(x)$  such that for every  $i \in N \setminus S$ ,  $u_i(x^k) < \varepsilon/n$  and each element of  $\pi(x^k)$  is either a winning coalition or a singleton. If  $x$  itself satisfies these properties, then we are done.



Otherwise, there is  $j \in N \setminus S$  such that  $u_j(x) \geq \varepsilon/n$ . Since  $j$  is not a veto player, it holds that  $N \setminus \{j\} \in \mathcal{W}$ . Let  $x^1 \in X$  be such that  $\pi(x^1) = \{N \setminus \{j\}, \{j\}\}$ ,  $u_{N \setminus \{j\}}(x^1) \gg u_{N \setminus \{j\}}(x)$ , and, for every  $i \in N \setminus \{j\}$  such that  $u_i(x) < \varepsilon/n$ ,  $u_i(x^1) < \varepsilon/n$ . Since  $u_j(x) \geq \varepsilon/n$ , such an element  $x^1$  exists. It holds that  $x^1 \in f(x)$  and  $u_j(x^1) = 0$ . If there is  $j^1 \in N \setminus S$  such that  $u_{j^1}(x^1) \geq \varepsilon/n$ , then we repeat this argument using  $j^1$ . Since the set  $N \setminus S$  is finite, we reach a state  $x^k$  with the desired properties in a finite number of steps.

We define the possibly empty set  $W^0 = \{i \in W \mid u_i(y) \leq \varepsilon/n\}$ . Let  $w \in W$  be a player such that  $u_w(y) \geq 1/n$ . Let  $z \in X$  be such that  $\pi(z) = \pi(y) = \{W\} \cup \{\{i\} \mid i \in N \setminus W\}$  and

$$u_j(z) = \begin{cases} \frac{\varepsilon}{n}, & j \in W^0, \\ u_j(y), & j \in W \setminus (W^0 \cup \{w\}), \\ u_w(y) - \sum_{i \in W^0} (\frac{\varepsilon}{n} - u_i(y)), & j = w. \end{cases}$$

For every  $j \in W^0$  it holds that  $u_j(z) = \varepsilon/n > u_j(x^k)$ , for every  $j \in W \setminus (W^0 \cup \{w\})$  it holds that  $u_j(z) = u_j(y) > \varepsilon/n > u_j(x^k)$ , and

$$u_w(z) = u_w(y) - \sum_{i \in W^0} (\frac{\varepsilon}{n} - u_i(y)) \geq \frac{1}{n} - \frac{n-2}{n} \frac{\varepsilon}{n} > \frac{\varepsilon}{n} > u_w(x^k),$$

so  $u_W(z) \gg u_W(x^k)$ . Since every element of  $\pi(x^k)$  is either a winning coalition or a singleton, it holds that  $W \in E(x^k, z)$ . We have shown that  $z \in f(x^k)$  and therefore  $z \in f^{k+1}(x)$ . We have that  $\pi(y) = \pi(z)$ , for every  $j \in W^0$  it holds that  $|u_j(y) - u_j(z)| \leq \varepsilon/n$ , for every  $j \in W \setminus (W^0 \cup \{w\})$  it holds that  $|u_j(y) - u_j(z)| = 0$ , and  $|u_w(y) - u_w(z)| \leq (n-2)\varepsilon/n$ , therefore  $z \in B_\varepsilon(y)$ , so  $z$  has all the desired properties.

It follows by Theorem 14 that  $F(X)$  is a subset of the MSS and since  $F(X)$  satisfies deterrence of external deviations and asymptotic external stability, it must be equal to the MSS.  $\square$

**PROOF OF THEOREM 31:** It is easy to see that the effectivity correspondence  $E$  is lower hemi-continuous. Moreover, continuity of  $(\geq_i)_{i \in N}$  is identical to continuity of the utility functions  $(u_i)_{i \in N}$ . As such, it suffices to show that  $\Gamma$  satisfies the weak improvement property.

Consider a function  $b : \Sigma \rightarrow \Sigma$  such that, for every  $s \in \Sigma$ ,

$$b(s) \in \arg \max_{\{\tau \in \Sigma \mid \exists i \in N, \tau_{-i} = s_{-i}\}} P(\tau),$$

and  $b(s) = s$  if  $P(b(s)) = P(s)$ . For each strategy profile  $s \in \Sigma$ , there is  $j \in N$  and  $\tau_j^* \in \Sigma_j$  such that  $b(s) = (\tau_j^*, s_{-j})$  and  $b(s)$  maximizes the value of  $P(\tau)$  over all strategies  $\tau$  such that there is  $i \in N$  with  $\tau_{-i} = s_{-i}$ . If the maximal value of  $P(\tau)$  is equal to  $P(s)$ , then  $b(s)$  is taken equal to  $s$ . As  $P$  is continuous and, for every  $i \in N$ ,  $\Sigma_i$  is compact, the maximization problem has a solution.

Observe that  $s \in \Sigma$  is a pure strategy Nash equilibrium of  $G$  if and only if

$$P(s) \geq P(\tau_i, s_{-i}), \quad i \in N, \tau_i \in \Sigma_i.$$

It follows that  $s$  is a pure strategy Nash equilibrium of  $G$  if and only if  $b(s) = s$ . Let NE be the set of all pure strategy Nash equilibria of  $G$ . The set NE is non-empty as it contains all the strategy profiles that maximize  $P$  over  $\Sigma$ , a non-empty set since  $P$  is continuous and  $\Sigma$  is compact.

Let us show that  $\Gamma$  satisfies the weak improvement property, i.e. for all states  $s \in \Sigma$ ,  $f^\infty(s) \cap \text{NE} \neq \emptyset$ .

Let  $s^1 \in \Sigma \setminus \text{NE}$  be given. Consider the sequence of states  $(s^k)_{k \in \mathbb{N}}$  in  $\Sigma$  defined by  $s^2 = b(s^1)$ ,  $s^3 = b(s^2)$ ,  $\dots$ . For every  $k \in \mathbb{N}$ , it holds that  $s^{k+1} \in f(s^k)$ , so any accumulation point of the sequence  $(s^k)_{k \in \mathbb{N}}$  belongs to  $f^\infty(s^1)$ .

Observe that, by definition,  $P(s^1) \leq P(s^2) \leq P(s^3) \leq \dots$ , so the sequence  $(P(s^k))_{k \in \mathbb{N}}$  is non-decreasing. Further, as  $(s^k)_{k \in \mathbb{N}}$  takes values in the compact set  $\Sigma$ , it has a convergent subsequence. Let us denote such a subsequence by  $(x^\ell)_{\ell \in \mathbb{N}}$  and let  $x^\ell \rightarrow x$ . It holds that  $x \in f^\infty(s^1)$ . By continuity of  $P$  it holds that  $P(x^\ell) \rightarrow P(x)$ . It also holds that, for every  $k \in \mathbb{N}$ ,  $P(s^k) \leq P(x)$ .

Since  $x \in f^\infty(s^1)$ , we finish the proof by showing that  $x \in \text{NE}$ . Suppose not, then there is  $j \in N$  and a best response  $\tau'_j \in \Sigma_j$  such that  $u_j(\tau'_j, x_{-j}) > u_j(x)$ . As  $G$  is a pseudo-potential game, there is a best response  $\tau_j^* \in \Sigma_j$  such that  $P(\tau_j^*, x_{-j}) > P(x)$ . Since  $P$  is continuous, there is  $\varepsilon > 0$  such that for every  $\tau \in B_\varepsilon((\tau_j^*, x_{-j}))$  it holds that  $P(\tau) > P(x)$ . As  $x^\ell \rightarrow x$ , there is  $\ell' \in \mathbb{N}$  such that  $x^{\ell'} \in B_\varepsilon(x)$ , so  $(\tau_j^*, x_{-j}^{\ell'}) \in B_\varepsilon((\tau_j^*, x_{-j}))$  and  $P(\tau_j^*, x_{-j}^{\ell'}) > P(x)$ . We have that

$$P(x) < P((\tau_j^*, x_{-j}^{\ell'})) \leq P(b(s^{\ell'})) \leq P(x),$$

a contradiction. Consequently, it holds that  $x \in \text{NE}$ .

## 2.7 Supplement

### 2.7.1 Examples

**Example 1.** In this example, we describe a social environment which has no MSS when iterated external stability is used in the definition instead of asymptotic external stability.

Consider the social environment  $\Gamma = (\{1\}, (X, d), E, \geq_1)$ , where the state space is given by  $X = \{1/k | k \in \mathbb{N}\} \cup \{0\}$  and  $d(x, y) = |x - y|$ . Note that  $X$  is compact. Preferences  $\geq_1$  are defined by  $x \geq_1 y$  if and only if  $x = y$  or  $x < y$ . The effectivity correspondence  $E$  is such that  $\{1\} \in E(1/k, 1/(k+1))$  for every  $k \in \mathbb{N}$  and  $E(x, y) = \emptyset$  otherwise. It follows that

$$f\left(\frac{1}{k}\right) = \left\{\frac{1}{k}, \frac{1}{k+1}\right\}.$$

Observe that  $0 \in f^\infty(x)$  for every  $x \in X$  and that  $f(0) = \{0\}$ . It now follows easily that  $\{0\}$  is an MSS.

We show that there is no closed set satisfying iterated external stability together with deterrence of external deviations and minimality. Towards a contradiction, suppose that the closed set  $M \subseteq X$  satisfies these properties. Since, for every  $k \in \mathbb{N}$ ,  $0 \notin f(1/k)$ , the set  $\{0\}$  does not satisfy iterated external stability. Given that  $M \neq \{0\}$  and  $M$  is non-empty, there is  $k \in \mathbb{N}$  such that  $1/k \in M$ . Let  $k$  be the smallest such number. From deterrence of external deviations, we have that also  $1/(k+1) \in M$ . Based on the corresponding properties of  $M$ , it is easy to verify that the closed, non-empty set  $M' = M \setminus \{1/k\}$  satisfies deterrence of external deviations and iterated external stability. Since  $M'$  is a proper subset of  $M$ ,  $M$  violates the minimality property.

**Example 2.** In the next example, we consider an infinite social environment for which there is more than one MSS.

Consider the social environment  $\Gamma = (\{1\}, (X, d), E, \geq_1)$ , where

$$X = \{0, \frac{1}{2}, 1\} \cup \{\frac{1}{k} \mid k \in \mathbb{N} \setminus \{1, 2\}\} \cup \{1 - \frac{1}{k} \mid k \in \mathbb{N} \setminus \{1, 2\}\}$$

and the metric is given by  $d(x, y) = |x - y|$ .

The effectivity correspondence is such that the individual can move from both states 0 and 1 to state  $1/2$  and, for every  $k \in \mathbb{N} \setminus \{1, 2\}$ , from state  $1 - 1/k$  to state  $1/k$  and from state  $1/k$  to state  $1 - 1/(k + 1)$ . The individual cannot make any other moves. The preferences of the individual are such that

$$\frac{2}{3} <_1 \frac{1}{3} <_1 \frac{3}{4} <_1 \frac{1}{4} <_1 \frac{4}{5} <_1 \frac{1}{5} <_1 \dots <_1 1 <_1 0 <_1 \frac{1}{2}.$$

We claim that both  $\{0, 1/2\}$  and  $\{1/2, 1\}$  are myopic stable sets. Since the effectivity correspondence admits no move outside the respective sets, both  $\{0, 1/2\}$  and  $\{1/2, 1\}$  satisfy deterrence of external deviations. For asymptotic external stability, observe that for every  $k \in \mathbb{N} \setminus \{1, 2\}$  it holds that  $\{0, 1\} \subset f^\infty(1/k)$  and  $\{0, 1\} \subset f^\infty(1 - 1/k)$ . Moreover, we have  $1/2 \in f(0) = f^\infty(0)$  and  $1/2 \in f(1) = f^\infty(1)$ . For minimality, the sets  $\{0\}$  and  $\{1\}$  violate deterrence of external deviations since  $1/2 \in f(0)$  and  $1/2 \in f(1)$ . The set  $\{1/2\}$  violates asymptotic external stability as  $1/2 \notin f^\infty(x)$  for any  $x \in X$  different from 0,  $1/2$  and 1.

**Example 3.** This example provides a social environment in which the effectivity correspondence is lower hemi-continuous and the preferences are continuous, but where the weak dominance MSS is not unique.

Consider the social environment  $\Gamma = (\{1, 2\}, (X, d), E, (\geq_1, \geq_2))$ , where

$$X = \{(0, 0), (1, 0), (2, 0)\} \cup \{(0, \frac{2}{k}), (1, \frac{1}{k}), (2, \frac{1}{k}) \mid k \in \mathbb{N}\}$$

and  $d$  is the Euclidean metric on  $X$ , so  $d(x, y) = \|x - y\|_2$ . It clearly holds that  $X$  is compact.

Individual 1 only cares about the first component of the state while individual 2 only cares about the second component. Both individuals prefer states where the component they care about is lower over states where it is higher. Note that these preferences are continuous.

The effectivity correspondence is as follows. For every  $k \in \mathbb{N}$ , the singleton  $\{1\}$  can move from state  $(2, 1/k)$  to state  $(1, 1/k)$  and the singleton  $\{2\}$  can move from state  $(1, 1/k)$  to state  $(2, 1/(k + 1))$ . Moreover, for every  $k \in \mathbb{N}$ , the singleton  $\{2\}$  can move from state  $(0, 2/k)$  to state  $(1, 1/k)$ . Coalition  $\{1, 2\}$  can move from states  $(1, 0)$  and  $(2, 0)$  to state  $(0, 0)$  and, for every  $k \in \mathbb{N}$ , from states  $(1, 1/k)$  and  $(2, 1/k)$  to state  $(0, 2/k)$ . No other moves are possible.

To see that the effectivity correspondence is lower hemi-continuous, let the sequence  $(x^k)_{k \in \mathbb{N}}$  in  $X$  be such that  $x^k \rightarrow x$ . There are only three relevant sequences of states in  $X$ : the sequence  $((0, 2/k))_{k \in \mathbb{N}}$ , the sequence  $((1, 1/k))_{k \in \mathbb{N}}$ , and the sequence  $((2, 1/k))_{k \in \mathbb{N}}$ . The first converges to  $(0, 0)$ , the second to  $(1, 0)$ , and the third to  $(2, 0)$ .

Let some  $x \in \{(0, 0), (1, 0), (2, 0)\}$  be given. Since  $G_{\{1\}}(x) = \{x\}$  and  $G_{\{2\}}(x) = \{x\}$ , it is immediate that  $G_{\{1\}}$  and  $G_{\{2\}}$  are lower hemi-continuous.

For  $G_{\{1, 2\}}$ , the only non-trivial cases are  $x = (1, 0)$  and  $x = (2, 0)$ . We give the argument for state  $x = (1, 0)$  explicitly. The argument for state  $(2, 0)$  follows by symmetry. For every  $y \in G_{\{1, 2\}}(1, 0)$  we have to find a sequence  $(y^k)_{k \in \mathbb{N}}$  such that  $y^k \in G_{\{1, 2\}}(1, 1/k)$  and

$y^k \rightarrow y$ . If  $y = (0, 0)$ , we take the sequence  $((0, 2/k))_{k \in \mathbb{N}}$ . If  $y = (1, 0)$ , we take the sequence  $((1, 1/k))_{k \in \mathbb{N}}$ .

Since  $f^\infty(0, 0) = \{(0, 0)\}$ ,  $f^\infty(1, 0) = \{(1, 0)\}$ , and  $f^\infty(2, 0) = \{(2, 0)\}$ , it follows from asymptotic external stability that  $\{(0, 0), (1, 0), (2, 0)\}$  is a subset of any MSS. Since this set satisfies deterrence of external deviations and asymptotic external stability, it follows from minimality that the unique MSS is equal to  $\{(0, 0), (1, 0), (2, 0)\}$ .

On the other hand, both sets  $\{(0, 0), (1, 0)\}$  and  $\{(0, 0), (2, 0)\}$  are a weak dominance MSS. Indeed, from both  $(1, 0)$  and  $(2, 0)$ , the coalition  $\{1, 2\}$  can deviate to  $(0, 0)$  if only weak dominance is imposed. To satisfy asymptotic external stability, it is sufficient that on top of the state  $(0, 0)$ , either the state  $(1, 0)$  or the state  $(2, 0)$  should be present. By minimality, it follows that only one of these states is included.

**Example 4.** This example demonstrates that the stochastic approach to infinite environments based on irreducibility of the Markov chain can deliver predictions that differ drastically from those of the MSS.

Consider the social environment  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$ , where  $N = \{1, 2\}$ ,  $X = [0, 1] \times [0, 1]$ , and the metric is  $d(x, y) = \|x - y\|_1 = |x_1 - y_1| + |x_2 - y_2|$ . The effectivity correspondence is such that individual 1 can change the first component of the state and individual 2 the second component, so  $\{1\} \in E(x, y)$  if and only if  $x_2 = y_2$  and  $\{2\} \in E(x, y)$  if and only if  $x_1 = y_1$ . The coalition  $\{1, 2\}$  is never effective. The preferences of the individuals are such that

$$\begin{aligned} x \geq_1 y & \text{ if and only if } 2x_1x_2 - x_1 - x_2 \geq 2y_1y_2 - y_1 - y_2, \\ x \geq_2 y & \text{ if and only if } 2x_1x_2 - x_1 - x_2 \leq 2y_1y_2 - y_1 - y_2. \end{aligned}$$

It is not hard to see that this social environment corresponds to the normal-form game of matching pennies, where  $x_1$  is the probability of the row player choosing "up" and  $x_2$  is the probability of the column player choosing "left". The unique Nash equilibrium of this game is equal to  $x^* = (1/2, 1/2)$ .

For every  $x \in X$ , we define

$$\begin{aligned} f_1(x) &= \begin{cases} \{y \in X \mid y_1 \leq x_1 \text{ and } y_2 = x_2\} & \text{if } x_2 < \frac{1}{2}, \\ \{x\} & \text{if } x_2 = \frac{1}{2}, \\ \{y \in X \mid y_1 \geq x_1 \text{ and } y_2 = x_2\} & \text{if } x_2 > \frac{1}{2}, \end{cases} \\ f_2(x) &= \begin{cases} \{y \in X \mid y_1 = x_1 \text{ and } y_2 \geq x_2\} & \text{if } x_1 < \frac{1}{2}, \\ \{x\} & \text{if } x_1 = \frac{1}{2}, \\ \{y \in X \mid y_1 = x_1 \text{ and } y_2 \leq x_2\} & \text{if } x_1 > \frac{1}{2}, \end{cases} \end{aligned}$$

so we can express the dominance correspondence as

$$f(x) = f_1(x) \cup f_2(x).$$

We consider the better-response dynamics where each element of  $f(x)$  is selected with equal probability. To do so, we define  $\rho_1 : X \rightarrow [0, 1]$  and  $\rho_2 : X \rightarrow [0, 1]$  as the functions that project  $x$  on its first and second coordinate, respectively. We use  $\lambda$  to denote the Lebesgue measure. Let  $\mathcal{B}(X)$  denote the Borel  $\sigma$ -algebra on  $X$ . The transition probability kernel resulting from better-response dynamics is obtained by defining, for every  $x \in X$ ,

and for every  $A \in \mathcal{B}(X)$ ,

$$Q(x, A) = \begin{cases} 0 & \text{if } x = (\frac{1}{2}, \frac{1}{2}) \notin A, \\ 1 & \text{if } x = (\frac{1}{2}, \frac{1}{2}) \in A, \\ 2\lambda(\rho_1(A \cap f_1(x))), & \text{if } x_1 = \frac{1}{2}, x_2 \neq \frac{1}{2}, \\ 2\lambda(\rho_2(A \cap f_2(x))), & \text{if } x_1 \neq \frac{1}{2}, x_2 = \frac{1}{2}, \\ \frac{\lambda(\rho_1(A \cap f_1(x))) + \lambda(\rho_2(A \cap f_2(x)))}{\lambda(\rho_1(f_1(x))) + \lambda(\rho_2(f_2(x)))}, & \text{if } x_1 \neq \frac{1}{2}, x_2 \neq \frac{1}{2}. \end{cases}$$

The first and second equality above show that the better-response dynamics never leaves the Nash equilibrium once reached. The third equality concerns the case where only player 1 likes to move. Observe that if  $x_1 = 1/2$  and  $x_2 \neq 1/2$ , then  $\lambda(\rho_1(f_1(x))) = 1/2$ , which explains the multiplication by 2. A similar remark applies to the fourth equality above. For the last equality, notice that  $x_1 \neq 1/2$  and  $x_2 \neq 1/2$  implies that  $\lambda(\rho_1(f_1(x))) > 0$  or  $\lambda(\rho_2(f_2(x))) > 0$ , so there is no division by zero.

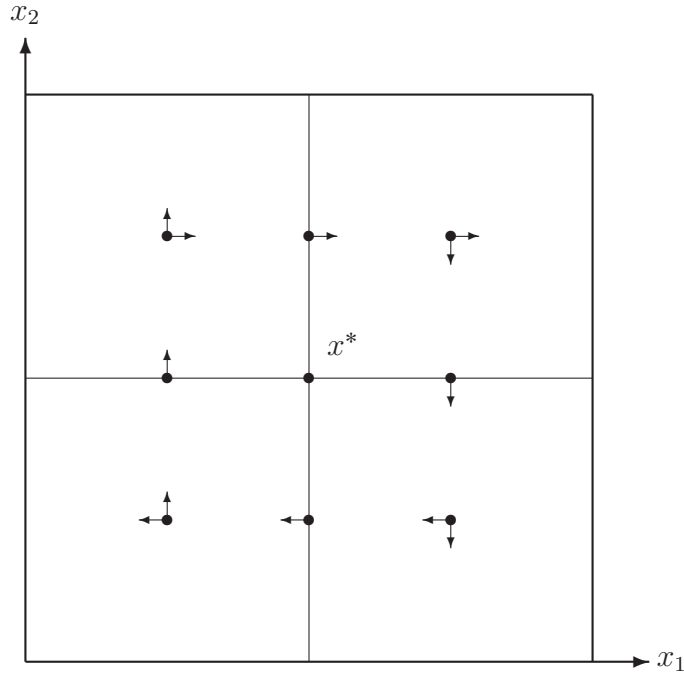


Figure 2.1: Better-response dynamics for the game of matching pennies.

The Markov process is illustrated in Figure 1. The arrows indicate the direction in which a state changes. A typical state can change in two directions, either west or east and either north or south, thereby generating two line segments on which the next state lies.

For every  $A \in \mathcal{B}(X)$ ,  $Q(\cdot, A)$  is a measurable function on  $X$ , but it is in general not continuous. For instance, if  $A = \{x^*\}$ , then  $Q(x, A) = 1$  if  $x = x^*$  and  $Q(x, A) = 0$ , otherwise. Indeed, the state  $x^*$  does not belong to  $f(x)$  unless  $x = x^*$  and in that case  $f(x^*) = \{x^*\}$ .

In this setting and other settings with an infinite state space, the Markov chain returns to a given state with probability zero, so the concept of a recurrent state is of less use and

importance. Instead, for infinite settings, the property of irreducibility is often studied, which expresses that all parts of the state space can be reached by the Markov chain, no matter what the starting point is. Given a state  $x \in X$  and a set  $A$  in the Borel  $\sigma$ -algebra  $\mathcal{B}(X)$  on  $X$ , let  $L(x, A)$  denote the probability that the Markov chain has a realization belonging to  $A$  at some point in the future when starting from  $x$ . Let  $\varphi$  be the measure on  $X$  that assigns to each set in  $\mathcal{B}(X)$  its Lebesgue measure. A Markov process  $(X, Q)$  is called  $\varphi$ -irreducible if for every  $A \in \mathcal{B}(X)$  such that  $\varphi(A) > 0$  it holds that  $L(x, A) > 0$  for every  $x \in X$ .

The Markov process  $(X, Q)$  in Example 4 is such that  $X$  can be decomposed in two parts, namely  $\{x^*\}$  and  $X \setminus \{x^*\}$ . There is no transition between these two sets of states and the restriction of the Markov process to each set is irreducible. This is obvious for  $\{x^*\}$ . The next result shows this for  $X \setminus \{x^*\}$ .

**Theorem 33.** The restriction of the Markov process  $(X, Q)$  in Example 4 to  $X \setminus \{x^*\}$  is  $\varphi$ -irreducible.

*Proof.* According to Proposition 4.2.1 of Meyn and Tweedie (1993), we have to show that for every  $x \in X \setminus \{x^*\}$ , for every  $A \in \mathcal{B}(X \setminus \{x^*\})$  such that  $\varphi(A) > 0$ , there exists  $k \in \mathbb{N}$  such that  $Q^k(x, A) > 0$ , where  $Q^k(x, A)$  denotes the probability of reaching  $A$  from  $x$  in  $k$  transitions.

It is convenient to partition the set  $X \setminus \{x^*\}$  in four subsets,

$$\begin{aligned} X^1 &= \{x \in X \mid x_1 \leq \frac{1}{2}, x_2 > \frac{1}{2}\}, \\ X^2 &= \{x \in X \mid x_1 > \frac{1}{2}, x_2 \geq \frac{1}{2}\}, \\ X^3 &= \{x \in X \mid x_1 \geq \frac{1}{2}, x_2 < \frac{1}{2}\}, \\ X^4 &= \{x \in X \mid x_1 < \frac{1}{2}, x_2 \leq \frac{1}{2}\}. \end{aligned}$$

Let some  $x \in X^4$  and some  $A \in \mathcal{B}(X \setminus \{x^*\})$  such that  $\varphi(A) > 0$  be given. We partition  $A$  in the four subsets  $A^1 \subseteq X^1$ ,  $A^2 \subseteq X^2$ ,  $A^3 \subseteq X^3$ , and  $A^4 \subseteq X^4$ . At least one of these four sets has positive Lebesgue measure. From  $x$ , the probability to reach a point in the set  $Y^1 = \{y^1 \in X^1 \mid y_1^1 = x_1\}$  is at least  $1/3$  and the probability distribution over  $Y^1$  is uniform. From  $y^1 \in Y^1$ , the probability to reach a point in the set  $Y^2(y^1) = \{y^2 \in X^2 \mid y_2^2 = y_1^1\}$  is at least  $1/3$  and the probability distribution over  $Y^2(y^1)$  is uniform. Thus, the probability to reach a point in  $X^2$  after 2 transitions is at least  $1/9$  and, conditional on reaching  $X^2$ , the distribution of this point is uniform on  $X^2$ . It now follows that  $Q^2(x, A) \geq \varphi(A^2)/9$ . Repeating this argument, we find that  $Q^3(x, A) \geq \varphi(A^3)/27$ ,  $Q^4(x, A) \geq \varphi(A^4)/81$ , and  $Q^5(x, A) \geq \varphi(A^1)/243$ . Since at least one of  $A^1$ ,  $A^2$ ,  $A^3$ , and  $A^4$  has strictly positive Lebesgue measure, we have shown that the restriction of the Markov process to  $X \setminus \{x^*\}$  is  $\varphi$ -irreducible. An analogous argument holds for  $x \in X^i$ , where  $i \neq 4$ .  $\square$

Example 4 shows that for the social environment corresponding to the normal-form game of matching pennies, none of the strategy profiles is singled out by the stochastic better-response dynamics. In contrast, we show in Section 2.7.6 that the MSS is unique and consists of the Nash equilibrium  $x^*$ .

**Example 5.** In this example, we show that the coalition structure core does not satisfy iterated external stability.

Let  $(N, v)$  be a coalition function form game such that  $N = \{1, 2, 3\}$ ,  $v(\{1, 2\}) = 1$ , and  $v(\{2, 3\}) = 1$ . All other coalitions have a coalitional value of 0. Thus, player 2 can choose to form a coalition with either player 1 or player 3 to form a two-person coalition generating a surplus equal to one. The coalition structure core therefore consists of only two states,  $y$  and  $y'$ , with equal payoffs,  $u(y) = u(y') = (0, 1, 0)$ , and coalitional structures  $\pi(y) = \{\{1, 2\}, \{3\}\}$  and  $\pi(y') = \{\{1\}, \{2, 3\}\}$ .

Consider an initial state  $x^0 \in X$  such that  $\pi(x^0) = \{\{1\}, \{2\}, \{3\}\}$  and  $u(x^0) = (0, 0, 0)$ . Under our notion of a myopic improvement, where all players involved in a move have to gain strictly, a state  $x^1 \neq x^0$  belongs to  $f(x^0)$  if and only if either  $\pi(x^1) = \{\{1, 2\}, \{3\}\}$  and  $u(x^1) = (\varepsilon, 1 - \varepsilon, 0)$  for some  $\varepsilon \in (0, 1)$  or  $\pi(x^1) = \{\{1\}, \{2, 3\}\}$  and  $u(x^1) = (0, 1 - \varepsilon, \varepsilon)$  for some  $\varepsilon \in (0, 1)$ . It follows that  $x^1$  is a state where either player 1 or player 3 receives a payoff of zero and the other two players receive a strictly positive payoff summing up to 1.

Now consider any state  $x^k$  such that either player 1 or player 3 receives 0 and the other two players receive a strictly positive payoff summing up to 1. We claim that any state  $x^{k+1} \in f(x^k)$  has the same properties. Without loss of generality, assume that  $u_3(x^k) = 0$ . Let  $x^{k+1}$  be an element of  $f(x^k)$  different from  $x^k$ . Since  $u_1(x^k) + u_2(x^k) = 1$ , the moving coalition is  $\{2, 3\}$  and it holds that  $\pi(x^{k+1}) = \{\{1\}, \{2, 3\}\}$ . Moreover, it must also hold that  $u_2(x^{k+1}) > u_2(x^k) > 0$  and  $u_3(x^{k+1}) > u_3(x^k) = 0$ , which proves the claim. Thus, for every  $k \in \mathbb{N}$ , if  $x^k \in f^k(x^0) \setminus \{x^0\}$ , then  $x^k$  is such that there are two players with a strictly positive payoff. It follows that there is no  $k \in \mathbb{N}$  such that  $x^k$  belongs to the coalition structure core.

## 2.7.2 Three-player Simple Games and the vNM Stable Set

Let  $(N, v)$  be a coalition function form game with  $N = \{1, 2, 3\}$  corresponding to a proper simple game. Let  $\Gamma$  be the social environment induced by the  $\gamma$ -model. We compare the prediction of the MSS with the vNM stable set for the three most interesting cases: there is one winning two-player coalition, without loss of generality  $\{1, 3\}$ ; there are two winning two-player coalitions, without loss of generality  $\{1, 2\}$  and  $\{2, 3\}$ ; all two-player coalitions are winning. The second case is known as the three-person veto-power game and the third case as the three-person simple majority game.

The first example shows that if  $\{1, 3\}$  is the only winning two-player coalition, then the MSS and the vNM stable set are unique and equal to the coalition structure core. The prediction is therefore that either coalition  $\{1, 3\}$  or coalition  $\{1, 2, 3\}$  forms and payoffs are such that the entire surplus is shared between players 1 and 3.

**Example 6.** Assume coalition  $\{1, 3\}$  is the only winning two-player coalition and singletons are not winning. By direct computation or by Step 2 of the proof of Theorem 25, it holds that the core of  $\Gamma$  is equal to the coalition structure core of  $(N, v)$ , so

$$Y = \{y \in X \mid \{1, 3\} \in \pi(y) \text{ or } [\{1, 2, 3\} \in \pi(y) \text{ and } u_2(y) = 0]\}.$$

By Theorem 25, the MSS of  $\Gamma$  is unique and equal to  $Y$ . So either coalition  $\{1, 3\}$  forms or the grand coalition forms and payoffs are such that the entire surplus is shared between players 1 and 3.

We argue that the vNM stable set is unique and equal to  $Y$  as well. Let  $V$  be a vNM stable set. For every  $y \in Y$  it holds that  $f(y) = \{y\}$ , so by external stability  $Y \subseteq V$ . We show

that  $Y$  satisfies external stability. Let  $x \notin Y$  be given. If  $\{1, 2, 3\} \notin \pi(x)$ , then  $x$  does not contain a winning coalition, so  $u(x) = (0, 0, 0)$ , and  $y \in Y$  defined by  $\pi(y) = \{\{1, 3\}, \{2\}\}$  and  $u(y) = (1/2, 0, 1/2)$  satisfies  $y \in f(x)$ . If  $\{1, 2, 3\} \in \pi(x)$ , then  $x \notin Y$  implies  $u_2(x) > 0$ . Now  $y \in Y$  defined by  $\pi(y) = \{\{1, 3\}, \{2\}\}$  and  $u(y) = (u_1(x) + u_2(x)/2, 0, u_3(x) + u_2(x)/2)$  satisfies  $y \in f(x)$ . We have shown that the core of  $\Gamma$  satisfies external stability. It must therefore be the unique vNM stable set.

We now turn to the three-person veto-power game, with player 2 being the veto player. The MSS is unique and equal to the coalition structure core, so one of the winning coalitions forms and player 2 gets the entire surplus. The MSS therefore has three elements, depending on the winning coalition that forms. We argue that there are two vNM stable sets, both having a continuum of elements and containing the MSS as a proper subset. In an element of the vNM stable set it holds that a winning coalition forms and the entire surplus is either shared between players 1 and 2 or between players 2 and 3. It is not excluded that the veto player gets a payoff of 0.

**Example 7.** Assume singletons are not winning and  $\{1, 2\}$  and  $\{2, 3\}$  are the winning two-player coalitions. By direct computation or by Step 2 of the proof of Theorem 25, it holds that the core of  $\Gamma$  is equal to the coalition structure core of  $(N, v)$ , so to the set

$$Y = \{y \in X \mid \pi(y) \cap \mathcal{W} \neq \emptyset \text{ and } u_2(y) = 1\}.$$

There are three states in  $Y$ . One of the winning coalitions  $\{1, 2\}$ ,  $\{2, 3\}$ , and  $\{1, 2, 3\}$  forms and players 1 and 3 receive a payoff of 0. By Theorem 25, the MSS of  $\Gamma$  is unique and equal to  $Y$ .

We argue that there are two vNM stable sets, both having a continuum of elements and containing the MSS as a proper subset. Let  $V$  be a vNM stable set. To satisfy external stability, it must hold that  $Y \subseteq V$ . Since states in  $Y$  do not dominate any other state, it follows by external stability that  $V$  contains  $Y$  as a proper subset. States  $x \in X$  such that  $\pi(x) \cap \mathcal{W} = \emptyset$  or  $\pi(x) = \{\{1, 2, 3\}\}$  do not dominate any state where a two-player winning coalition forms. It therefore follows from external stability that  $V$  contains a state  $x^1 \in X \setminus Y$  such that  $\{1, 2\} \in \pi(x^1)$  or  $\{2, 3\} \in \pi(x^1)$ . Without loss of generality, assume that  $\{1, 2\} \in \pi(x^1)$ . Notice that  $u_1(x^1) > 0$  since  $x^1 \in X \setminus Y$ . We distinguish between two cases: Case 1. There is  $x^2 \in V \setminus Y$  such that  $\{2, 3\} \in \pi(x^2)$ . Case 2. For every  $x \in V \setminus Y$  it holds that  $\{2, 3\} \notin \pi(x)$ .

Case 1. Since  $x^2 \in V \setminus Y$ , it holds that  $u_3(x^2) > 0$ . In order to satisfy internal stability, it must hold that  $u_2(x^1) = u_2(x^2)$  and therefore  $u_1(x^1) = u_3(x^2)$ . Internal stability implies that there cannot be any other  $x \in V \setminus Y$  such that  $\pi(x)$  contains  $\{1, 2\}$  or  $\{2, 3\}$ . Now  $x^3 \in X \setminus V$  such that  $\pi(x^3) = \{\{1, 2\}, \{3\}\}$  and  $u_2(x^1) < u_2(x^3)$  is not dominated by an element of  $V$ , so  $V$  does not satisfy external stability, and we have obtained a contradiction.

Case 2. None of the states  $x \in X$  such that  $\pi(x) = \{\{1, 2\}, \{3\}\}$  is dominated by a state in  $V$ , so every such state must belong to  $V$  to satisfy external stability. The same applies to a state  $x \in X$  such that  $\pi(x) = \{\{1, 2, 3\}\}$  and  $u_3(x) = 0$ . We have that the set

$$V' = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}$$

is a subset of  $V$ . Notice that  $V'$  contains a single element with  $\{2, 3\}$  as the winning coalition, a continuum of elements with  $\{1, 2\}$  as the winning coalition, and a continuum of



elements with  $\{1, 2, 3\}$  as the winning coalition. It is easily verified that all states in  $X \setminus V'$  are dominated by an element that belongs to  $V'$ . None of the elements in  $V'$  dominate each other. We have therefore shown that

$$V = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}.$$

By symmetry it follows that

$$\{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_1(x) = 0\}$$

is a vNM stable set as well. This exhausts all possibilities.

We finally turn to the three-player simple majority game. The MSS is unique and equal to the set of states such that a two-player winning coalition forms. We argue that there are four vNM stable sets, none of them being a subset of the MSS or containing the MSS as a subset. Every vNM stable set contains elements where the grand coalition forms. The union of the four vNM stable sets contains the MSS as a proper subset.

**Example 8.** Assume all two-player coalitions are winning, whereas all singletons are not winning. It follows from Theorem 26 that the MSS is unique and equal to the set

$$F(X) = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}\} \neq \emptyset\}.$$

In particular, it is excluded that the grand coalition forms.

Let  $V$  be a vNM stable set. Let  $x \in X$  be a state such that  $\pi(x) \cap \mathcal{W} = \emptyset$ , or  $\pi(x)$  contains a two-player winning coalition with one of the players in that coalition having a payoff of 1, or  $\pi(x) = \{\{1, 2, 3\}\}$ . Since  $x$  does not dominate any state in  $F(X)$ , it follows by external stability that  $V$  contains an element  $x^1 \in F(X)$  with payoffs being strictly positive for both players in the winning coalition. Without loss of generality, assume that  $\{1, 2\} \in \pi(x^1)$ . We distinguish between two cases: Case 1. There is  $x^2 \in V$  such that  $\{1, 3\}$  or  $\{2, 3\}$  belongs to  $\pi(x^2)$  and  $u_3(x^2) > 0$ . Case 2. For every  $x \in V$  it holds that if  $\{1, 3\}$  or  $\{2, 3\}$  belongs to  $\pi(x)$ , then  $u_3(x) = 0$ .

Case 1. Without loss of generality, assume  $\{1, 3\} \in \pi(x^2)$ . To satisfy internal stability, it must hold that  $u_1(x^1) = u_1(x^2)$  and therefore  $u_2(x^1) = u_3(x^2)$ . There cannot be a state  $x \in V \setminus \{x^1, x^2\}$  such that  $\{1, 2\} \in \pi(x)$  or  $\{1, 3\} \in \pi(x)$  and  $u_1(x) < 1$  since otherwise internal stability would be violated.

Suppose, in order to derive a contradiction, that there is no state  $x \in V$  with  $\{2, 3\} \in \pi(x)$ . If  $u_1(x^1) \geq 1/2$ , then the state  $x \in X$  such that  $\pi(x) = \{\{2, 3\}, \{1\}\}$  and  $u(x) = (0, 1/2, 1/2)$  is not dominated by an element of  $V$ . If  $u_1(x^1) < 1/2$ , then the state  $x \in X$  such that  $\pi(x) = \{\{1, 2\}, \{3\}\}$  and  $u(x) = (1/2, 1/2, 0)$  is not dominated by an element of  $V$ . Since  $V$  satisfies external stability, we have obtained a contradiction. Consequently, there is a state  $x^3 \in V$  such that  $\{2, 3\} \in \pi(x)$ .

In order not to violate internal stability, it must hold that  $u_2(x^3) = u_2(x^1)$  and  $u_3(x^3) = u_3(x^2)$ . Since  $u_2(x^1) = u_3(x^2)$ , this is only possible if  $u(x^3) = (0, 1/2, 1/2)$ . It follows that  $u(x^1) = (1/2, 1/2, 0)$  and  $u(x^2) = (1/2, 0, 1/2)$ . We define  $x^4, x^5, x^6 \in X$  by  $\pi(x^4) = \pi(x^5) = \pi(x^6) = \{\{1, 2, 3\}\}$  and  $u(x^4) = u(x^1)$ ,  $u(x^5) = u(x^2)$ , and  $u(x^6) = u(x^3)$ . It is easily verified that all states in  $X \setminus \{x^1, \dots, x^6\}$  are dominated by  $x^1, x^2$ , or  $x^3$ . The states  $x^1, \dots, x^6$  do not dominate each other. This yields  $V = \{x^1, \dots, x^6\}$  as the unique vNM stable set satisfying the assumptions of Case 1.

Case 2. None of the states  $x \in X$  such that  $\pi(x) = \{\{1, 2\}, \{3\}\}$  is dominated by a state in  $V$ , so every such state must belong to  $V$  to satisfy external stability. The same applies to states  $x \in X$  such that  $\pi(x) = \{\{1, 3\}, \{2\}\}$ ,  $\pi(x) = \{\{2, 3\}, \{1\}\}$ , or  $\pi(x) = \{\{1, 2, 3\}$  and  $u_3(x) = 0$ . We have that the set

$$V' = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}$$

is a subset of  $V$ . It is easily verified that all states in  $X \setminus V'$  are dominated by an element that belongs to  $V'$ . It follows that

$$V = \{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_3(x) = 0\}.$$

We can easily check that  $V$  satisfies internal stability as well.

By symmetry it follows that

$$\begin{aligned} &\{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_1(x) = 0\}, \\ &\{x \in X \mid \pi(x) \cap \{\{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\} \neq \emptyset \text{ and } u_2(x) = 0\}, \end{aligned}$$

are vNM stable sets as well. This exhausts all possibilities.

### 2.7.3 Proper Simple Games and the $\delta$ -Model

Let  $(N, v)$  be a coalition function form game. The effectivity correspondence  $E$  is said to be induced by the  $\delta$ -model if it satisfies coalitional sovereignty and for every  $x, y \in X$ , for every  $S \in E(x, y)$ , for every  $T \in \pi(x)$  such that  $T \setminus S \neq \emptyset$ , it holds that  $T \setminus S \in \pi(y)$ . The latter condition simply expresses that residual players in some coalition stay together after coalition  $S$  leaves. Typically, it is assumed that the change in payoffs of the residual players in a given coalition has the same sign. When we restrict the analysis to proper simple games, we can obtain a characterization of the MSS without any such additional assumptions.

Let  $(N, v)$  be a coalition function form game such that  $v$  is a proper simple game with an empty core. We define the subset  $F'(X)$  of  $X$  as the set of states such that its partition contains a winning coalition different from the grand coalition:

$$F'(X) = \{x \in X \mid \pi(x) \cap (\mathcal{W} \setminus \{N\}) \neq \emptyset\}.$$

It holds that  $F(X) \subseteq F'(X)$ , where  $F(X)$  is defined in the main text. The only difference between these two sets is that  $F'(X)$  does not require the non-winning coalitions in  $\pi(x)$  to be singletons. In terms of payoff vectors that can be supported, there is no difference between  $F(X)$  and  $F'(X)$ .

**Theorem 34.** Let  $(N, v)$  be a coalition function form game such that  $v$  is a proper simple game with an empty core and let  $\Gamma = (N, (X, d), E, (\succeq_i)_{i \in N})$  be the social environment induced by the  $\delta$ -model. It holds that the MSS of  $\Gamma$  is unique and equal to  $F'(X)$ .

*Proof.* The proof that  $F'(X)$  satisfies deterrence of external deviations follows exactly the same steps as the corresponding part of the proof of Theorem 26. The same is true for the proof of asymptotic external stability of  $F'(X)$ . The only difference is that there is no need to verify that the non-winning coalitions are singletons.

We complete the proof by showing that for every  $x \in F'(X)$  it holds that  $f^\infty(x) = F'(X)$ . Let some  $x, y \in F'(X)$  be given and denote the winning coalition in  $\pi(y)$  by  $W$ . We have to show that for every  $\varepsilon > 0$  there exists  $k' \in \mathbb{N}$  and  $z \in f^{k'}(x)$  such that  $z \in B_\varepsilon(y)$ .

Let some  $\varepsilon \in (0, 1/n)$  and  $S \in \pi(y) \setminus \{W\}$  be given. By following exactly the same steps as in the proof of Theorem 26 it can be shown that there is  $k \in \mathbb{N}$  and  $x^k \in f^k(x)$  such that for every  $i \in N \setminus S$ ,  $u_i(x^k) < \varepsilon/n$  and  $\pi(x^k)$  contains a winning coalition. Since  $\sum_{i \in N \setminus S} u_i(x^k) < 1$ , it follows that  $\sum_{i \in S} u_i(x^k) > 0$  and that  $S$  has a non-empty intersection with the winning coalition in  $\pi(x^k)$ .

Write  $\pi(y)$  as  $\{S^1, \dots, S^{\ell'}\}$  with  $S^1 = S$  and  $S^{\ell'} = W$ . For  $\ell = 1, \dots, \ell' - 2$ , let  $x^{k+\ell} \in X$  be such that  $\pi(x^{k+\ell}) = \{S^1, \dots, S^\ell, S^{\ell+1} \cup \dots \cup S^{\ell'}\}$ ,  $u_{S^{\ell+1} \cup \dots \cup S^{\ell'}}(x^{k+\ell}) \gg u_{S^{\ell+1} \cup \dots \cup S^{\ell'}}(x^{k+\ell-1})$  and for every  $i \in W$ ,  $u_i(x^{k+\ell}) < \varepsilon/n$ . In step  $\ell$ , coalition  $S^{\ell+1} \cup \dots \cup S^{\ell'}$  forms and increases the payoffs of its members, whereas the payoffs of the players in  $W$  are kept strictly below  $\varepsilon/n$ . Coalition  $S^\ell$  becomes part of  $\pi(x^{k+\ell})$  as a residual set of players. Since  $\sum_{i \in S^\ell} u_i(x^{k+\ell-1}) > 0$ , such a state  $x^{k+\ell}$  exists.

We define the possibly empty set  $W^0 = \{i \in W \mid u_i(y) \leq \varepsilon/n\}$ . Let  $w \in W$  be a player such that  $u_w(y) \geq 1/n$ . Let  $z \in X$  be such that  $\pi(z) = \pi(y) = \{S^1, \dots, S^{\ell'}\}$  and

$$u_j(z) = \begin{cases} \frac{\varepsilon}{n}, & j \in W^0, \\ u_j(y), & j \in W \setminus (W^0 \cup \{w\}), \\ u_w(y) - \sum_{i \in W^0} (\frac{\varepsilon}{n} - u_i(y)), & j = w. \end{cases}$$

For every  $j \in W^0$  it holds that  $u_j(z) = \varepsilon/n > u_j(x^{k+\ell'-2})$ , for every  $j \in W \setminus (W^0 \cup \{w\})$  it holds that  $u_j(z) = u_j(y) > \varepsilon/n > u_j(x^{k+\ell'-2})$ , and

$$u_w(z) = u_w(y) - \sum_{i \in W^0} (\frac{\varepsilon}{n} - u_i(y)) \geq \frac{1}{n} - \frac{n-2}{n} \frac{\varepsilon}{n} > \frac{\varepsilon}{n} > u_w(x^{k+\ell'-2}),$$

so  $u_W(z) \gg u_W(x^{k+\ell'-2})$ . It follows that  $z \in f(x^{k+\ell'-2})$  and therefore  $z \in f^{k+\ell'-1}(x)$ . We have that  $\pi(y) = \pi(z)$ , for every  $j \in W^0$  it holds that  $|u_j(y) - u_j(z)| \leq \varepsilon/n$ , for every  $j \in W \setminus (W^0 \cup \{w\})$  it holds that  $|u_j(y) - u_j(z)| = 0$ , and  $|u_w(y) - u_w(z)| \leq (n-2)\varepsilon/n$ , therefore  $z \in B_\varepsilon(y)$ , so  $z$  has all the desired properties.

It follows by Theorem 14 that  $F'(X)$  is a subset of the MSS and since  $F'(X)$  satisfies deterrence of external deviations and asymptotic external stability, it must be equal to the MSS.  $\square$

We now turn to the case where  $(N, v)$  is a proper simple game with a non-empty core and show that the analogue of Theorem 25 for the  $\delta$ -model holds.

**Theorem 35.** Let  $(N, v)$  be a coalition function form game such that  $v$  is a proper simple game with a non-empty core and let  $\Gamma = (N, (X, d), E, (\geq_i)_{i \in N})$  be the social environment induced by the  $\delta$ -model. It holds that the MSS of  $\Gamma$  is unique and equal to the coalition structure core  $Y$  of  $(N, v)$ .

*Proof.* The coalition structure core  $Y$  is the set of states such that a winning coalition forms, the other players are partitioned in arbitrary coalitions, and the veto players are the only ones with a positive payoff. Step 2 in the proof of Theorem 25 can be used to show that also for the  $\delta$ -model the core CO of  $\Gamma$  is equal to  $Y$ . Since CO is closed, the remark below Theorem 18 implies that we only have to show that  $\Gamma$  satisfies the weak improvement property.

We need to show that for every  $x \in X$ ,  $f^\infty(x) \cap Y \neq \emptyset$ . This is trivial if  $x \in Y$ . Assume  $x \in X \setminus Y$ . We have to show that for every  $\varepsilon > 0$  there exists  $k \in \mathbb{N}$ ,  $z \in f^k(x)$ , and  $y \in Y$  such that  $z \in B_\varepsilon(y)$ . Let some  $\varepsilon \in (0, 1/n)$  be given. It holds that either  $\pi(x) \cap \mathcal{W} = \emptyset$  or there is  $i' \in N \setminus S^*$  such that  $u_{i'}(x) > 0$ .

If  $\pi(x) \cap \mathcal{W} = \emptyset$ , then choose a winning coalition  $W \in \mathcal{W}$  and a veto player  $w \in W \cap S^*$ . Let  $z \in X$  be such that  $W \in E(x, z)$  and

$$\begin{aligned} z_j &= 0, & j &\in N \setminus W, \\ z_j &= \frac{\varepsilon}{n}, & j &\in W \setminus \{w\}, \\ z_j &= 1 - \sum_{i \in W \setminus \{w\}} \frac{\varepsilon}{n}, & j &= w. \end{aligned}$$

It holds that  $z \in f(x)$ . Let  $y \in Y$  be such that  $\pi(y) = \pi(z)$  and  $u_w(y) = 1$ . It holds that  $z \in B_\varepsilon(y)$ . This shows that  $z$  has the desired properties.

If there is  $i' \in N \setminus S^*$  such that  $u_{i'}(x) > 0$ , then let  $W$  be the unique element in  $\pi(x) \cap \mathcal{W}$ . We show first that there exists  $k \in \mathbb{N}$  and  $z \in f^k(x)$  such that, for every  $i \in N \setminus S^*$ ,  $u_i(z) < \varepsilon$ . If for every  $j \in W \setminus S^*$  it holds that  $u_j(x) < \varepsilon$ , then take  $z = x$ . Otherwise, there is  $j \in W \setminus S^*$  such that  $u_j(x) \geq \varepsilon$ . Since  $j$  is not a veto player, it holds that  $N \setminus \{j\} \in \mathcal{W}$ . Let  $x^1 \in X$  be such that  $\pi(x^1) = \{N \setminus \{j\}, \{j\}\}$ ,  $u_{N \setminus \{j\}}(x^1) \gg u_{N \setminus \{j\}}(x)$ , and, for every  $i \in N \setminus \{j\}$  such that  $u_i(x) < \varepsilon$ , it holds that  $u_i(x^1) < \varepsilon$ . Since  $u_j(x) \geq \varepsilon$ , such an element  $x^1$  exists. It holds that  $x^1 \in f(x)$  and  $u_j(x^1) = 0$ . If there is  $j^1 \in W \setminus S^*$  such that  $u_{j^1}(x^1) \geq \varepsilon$ , then we repeat this argument using  $j^1$ . Since the set  $W \setminus S^*$  is finite, we reach a state  $z$  with the desired properties in a finite number of steps. Clearly, there is  $y \in Y$  with  $\pi(y) = \pi(z)$  and  $z \in B_\varepsilon(y)$ .  $\square$

## 2.7.4 The vNM Stable Set for the Tamura Example of the Knuth Model

Let us reconsider the graph on page 316 of Tamura (1993). There is a total of 24 matchings, denoted by  $M_1, \dots, M_{24}$ . The core of the social environment induced by the Knuth (1976) model is equal to  $\text{CO} = \{M_1, M_8, M_{10}, M_{19}, M_{24}\}$ .

The MSS contains 13 matchings. In addition to the matchings in the core, we obtain 8 matchings in a closed cycle and find that the MSS is equal to

$$\text{CO} \cup \{M_2, M_{16}, M_{22}, M_{12}, M_7, M_9, M_3, M_4\}.$$

There are two different vNM stable sets in this example. The first vNM stable set  $V_1$  is given by

$$V_1 = \text{CO} \cup \{M_4, M_5, M_9, M_{12}, M_{13}, M_{16}, M_{17}, M_{20}, M_{21}\}.$$

Another vNM stable set is equal to

$$V_2 = \text{CO} \cup \{M_2, M_3, M_7, M_{11}, M_{14}, M_{17}, M_{18}, M_{22}, M_{23}\}.$$

The prediction of the vNM stable sets seems rather unappealing. First, the dominated state  $M_{17}$  is part of each vNM stable set. Second, for the largest connected subgraph of the divorce digraph, half of the states is in  $V_1$ , while the other half is in  $V_2$ .

agents	1	2	3
first choice	2	3	1
second choice	3	1	2
third choice	1	2	3

Table 2.1: A Shapley-Scarf housing matching market with cycling.

## 2.7.5 Shapley-Scarf Housing Markets

Another prominent matching model is the housing matching model of Shapley and Scarf (1974). This model can be represented by a tuple  $(N, H, (P_i)_{i \in N})$ , where  $N$  is a finite set of individuals,  $H$  is a finite set of houses with the same cardinality as the set of individuals, and each individual  $i \in N$  has a strict preference relation  $P_i$  over  $H$ . The original chapter by Shapley and Scarf (1974) does not require a strict preference relation. However, as shown in Roth and Postlewaite (1977), when preferences are strict, then the strong core, i.e., the core based on weak dominance, contains a unique element. The version with strict preferences therefore became popular in the literature. Without loss of generality, we assume that  $N = H$  and that the initial endowment of individual  $i$  is house  $i$ . An allocation is represented by a permutation matrix  $A$  with rows indexed by elements of  $N$  and columns indexed by elements of  $H$ . All entries of  $A$  are 0 or 1 and both rows and columns of  $A$  sum up to 1. If for some  $h \in H$ , for some  $i \in N$ , entry  $A_{ih} = 1$ , then house  $h$  has been assigned to individual  $i$ . Row  $i \in N$  of the matrix  $A$  is denoted by  $A_i$ .

In this setting, it is convenient to define the state space  $X$  as the set of all permutation matrices  $A$ . Since  $X$  is finite, we can endow it with the discrete metric  $d(A, A') = 1_{\{A \neq A'\}}$ .

The preferences of the individuals  $(\succeq_i)_{i \in N}$  over the set  $X$  are induced by their preferences over houses in the following way. Let some individual  $i \in N$  be given as well as  $A, A' \in X$ . Let  $h, h' \in H$  be such that  $A_{ih} = A'_{ih'} = 1$ . Notice that  $h$  and  $h'$  are uniquely determined. It holds that  $A \succ_i A'$  if and only if  $h P_i h'$ .

A coalition  $S \in \mathcal{N}$  can arbitrarily redistribute the initial endowments of houses of its members within the coalition. More formally, the effectivity correspondence satisfies the following two conditions:

1. For every  $S \in \mathcal{N}$ , for every  $A, A' \in X$ , if  $S \in E(A, A')$  then for all  $i \in S$ , there is  $h \in S$  such that  $A'_{ih} = 1$ .
2. For every  $S \in \mathcal{N}$ , for every  $A \in X$ , and for every bijection  $\phi : S \rightarrow S$ , there exists  $A' \in X$  such that for all  $i \in S$ ,  $A'_{i\phi(i)} = 1$  and  $S \in E(A, A')$ .

The first condition requires that if  $S$  is effective in moving from state  $A$  to state  $A'$ , then at  $A'$  the initial endowments of members of  $S$  are reallocated within  $S$ . The second condition states that every reallocation of initial endowments of houses within a coalition is feasible. Observe that the conditions impose no restrictions on how the houses of members outside the deviating coalition are reallocated, so we allow for various reallocation processes here. This completes the description of the social environment.

We show first that the MSS may contain closed cycles that do not correspond to a core element. We consider the example illustrated in Table 1. Since the initial endowments cor-

respond to every individual's worst choice, every allocation in  $X$  is individually rational. We argue that the set  $\{A^1, A^2, A^3\}$  is a closed cycle, where

$$A^1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, A^2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, A^3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

At  $A^1$ , individual 1 obtains his first best house. Since  $A^1$  is individually rational, coalition  $\{2, 3\}$  is the only coalition that can achieve a strict improvement. The only state that dominates  $A^1$  is therefore  $A^2$ . By the same argument, the only state that dominates  $A^2$  is  $A^3$ , and the only state that dominates  $A^3$  is  $A^1$ . We have obtained a closed cycle that does not contain a core element. By Theorem 14, the core of the housing market model is a proper subset of the MSS.

We now turn to the weak dominance MSS and show that it is equal to the strong core by using the top trading cycle algorithm of Shapley and Scarf (1974).

**Theorem 36.** Let  $(N, H, (P_i)_{i \in N})$  be a housing matching problem and let  $\Gamma$  be the induced social environment. The weak dominance MSS of  $\Gamma$  is equal to the strong core.

*Proof.* Since the strong core of the housing matching problem is unique and satisfies deterrence of external deviations, we only have to show that it satisfies iterated external stability. Let  $S^1, \dots, S^{k'}$  be the coalitions that are successively formed by an application of the top trading cycle algorithm of Shapley and Scarf (1974).

Consider any allocation  $A \in X$  that is not equal to the strong core allocation  $A^*$  of the housing matching model. Let  $\tilde{f}$  denote the weak dominance correspondence. We generate a sequence of allocations  $A^1, A^2, \dots, A^{k'}$  such that for all  $k \leq k'$ ,  $A^k \in \tilde{f}^k(A)$  and  $A^{k'} = A^*$ . Let  $A^0 = A$  and  $k = 1$ . We construct the sequence in the following way.

1. If  $k = k' + 1$  stop.
2. If for every  $i \in S^k$  it holds that  $A_i^{k-1} = A_i^*$ , then we set  $A^k = A^{k-1}$ . Increase  $k$  by one and go back to step 1.

If there is  $i \in S^k$  such that  $A_i^{k-1} \neq A_i^*$ , then define  $T^k = \cup_{j \leq k} S^j$ , and let  $A^k$  be an allocation such that  $T^k \in E(A^{k-1}, A^k)$  and, for every  $i \in T^k$ ,  $A_i^k = A_i^*$ . Increase  $k$  by one and go back to step 1.

We argue that for every  $k = 1, \dots, k'$ ,  $A^k \in \tilde{f}(A^{k-1})$ . This is trivial if for every  $i \in S^k$  it holds that  $A_i^{k-1} = A_i^*$ , since then  $A^k = A^{k-1}$ .

Let  $k \in \{1, \dots, k'\}$  and  $i \in S^k$  be such that  $A_i^{k-1} \neq A_i^*$ . By the rules of the top trading cycle algorithm, the house corresponding to  $A_i^*$  is the best house for  $i$  in the set of houses  $N \setminus T^{k-1}$ , so in particular it holds that  $A_i^* >_i A_i^{k-1}$ . It now follows that  $A^k \in \tilde{f}(A^{k-1})$ .

The proof is completed by observing that  $A^{k'} = A^*$ . □

In the proof of Theorem 36, we have the union  $T^k$  of the coalitions  $S^1, \dots, S^k$  generated in the first  $k$  steps of the top trading cycle algorithm deviating in iteration  $k$  of our construction. The reason is that the assumptions on the effectivity correspondence are so weak that a deviation by  $S^k$  might upset the assignment of individuals in  $T^{k-1}$ . Under stronger assumptions on the effectivity correspondence, like those corresponding to the  $\gamma$  or the  $\delta$ -model, it would be sufficient to have deviations by  $S^k$  in iteration  $k$ .

## 2.7.6 Mixed Environments

Let  $G = (N, ((\Sigma_i, d_i), u_i)_{i \in N})$  be a finite normal-form game, so for each player  $i \in N$  it holds that  $\Sigma_i$  is finite and  $d_i(s_i, s'_i) = 1_{\{s_i \neq s'_i\}}$ .

Let us now introduce the mixed extension  $\tilde{G} = (N, ((\Delta_i, \delta_i), v_i)_{i \in N})$  of  $G$ , where  $\Delta_i$  is the set of probability distributions on  $\Sigma_i$ . For  $\sigma_i \in \Delta_i$ ,  $\sigma_{i,s_i}$  denotes the probability that player  $i$  uses pure strategy  $s_i$ . The metric  $\delta_i$  on  $\Delta_i$  is defined by

$$\delta_i(\sigma_i, \sigma'_i) = \max_{s_i \in \Sigma_i} |\sigma_{i,s_i} - \sigma'_{i,s_i}|.$$

We denote  $\Delta = \prod_{i \in N} \Delta_i$  and endow  $\Delta$  with the product metric  $\delta(\sigma, \sigma') = \sum_{i \in N} \delta_i(\sigma_i, \sigma'_i)$ . For a given strategy profile  $\sigma \in \Delta$ , we denote the probability that pure strategy profile  $s \in \Sigma$  is played by  $\sigma_s = \prod_{i \in N} \sigma_{i,s_i}$ . Let  $v_i : \Delta \rightarrow \mathbb{R}$  be the expected utility associated to strategy profiles  $\sigma \in \Delta$ ,

$$v_i(\sigma) = \sum_{s \in \Sigma} \sigma_s u_i(s).$$

Preferences  $(\geq_i)_{i \in N}$  are such that  $\sigma \geq_i \sigma'$  if and only if  $v_i(\sigma) \geq v_i(\sigma')$ . The social environment  $\tilde{\Gamma} = (N, (\Delta, \delta), E, (\geq_i)_{i \in N})$  corresponds to the game  $\tilde{G}$  where  $E$  only allows singletons to deviate and  $\{i\} \in E(\sigma, \sigma')$  if and only if  $\sigma_{-i} = \sigma'_{-i}$ .

A strategy profile  $\sigma \in \Delta$  is said to be a mixed strategy Nash equilibrium of  $G$  if it is a pure strategy Nash equilibrium of  $\tilde{G}$ . The core of  $\tilde{\Gamma}$  coincides with the set of mixed strategy Nash equilibria of  $G$ . Additionally, note that the expected utility functions  $(v_i)_{i \in N}$  are continuous on  $\Delta$  and that  $E$  is lower hemi-continuous. As such, Theorems 12 and 18 give the following result.

**Corollary 37.** Let  $\tilde{G}$  be the mixed extension of the finite normal-form game  $G$  and let  $\tilde{\Gamma}$  be the social environment corresponding to  $\tilde{G}$ . The MSS of  $\tilde{\Gamma}$  coincides with the set of mixed strategy Nash equilibria of  $G$  if and only if  $\tilde{\Gamma}$  satisfies the weak improvement property.

Clearly, the pure strategy Nash equilibria of  $G$  are also mixed strategy Nash equilibria of  $G$ , so belong to the MSS of  $\tilde{\Gamma}$ . On the other hand, it is easy to find examples such that some profiles in the MSS of  $\tilde{\Gamma}$  are not in the MSS of  $\tilde{\Gamma}$ .

A finite two-player game  $G = (N, ((\Sigma_i, d_i), u_i)_{i \in \{1,2\}})$  is zero-sum if for all strategy profiles  $s \in \Sigma$ ,  $u_1(s) + u_2(s) = 0$ . The following result shows that for such games the MSS of  $\tilde{\Gamma}$  coincides with the set of mixed strategy Nash equilibria of  $G$ .

**Theorem 38.** Let  $\tilde{G}$  be the mixed extension of a finite two-player zero-sum game  $G$  and let  $\tilde{\Gamma}$  be the social environment corresponding to  $\tilde{G}$ . Then the MSS of  $\tilde{\Gamma}$  coincides with the set of mixed strategy Nash equilibria of  $G$ .

*Proof.* Using Corollary 37, it remains to show that  $\tilde{\Gamma}$  satisfies the weak improvement property, i.e., for every strategy profile  $\sigma \in \Delta$ ,  $f^\infty(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ . Let  $v$  denote the value of the game.

Let some  $\sigma \in \Delta$  be given which is not a mixed strategy Nash equilibrium of  $G$ , i.e., there is a player  $i$  such that  $\sigma_i$  is not a minmax strategy. We distinguish between two cases.

**Case 1:**  $\sigma_1$  and  $\sigma_2$  are not minmax strategies.

- 1.1 If  $v_1(\sigma) \neq v$ , then there exists a player  $i$  who is below his minmax payoff. Without loss of generality, let this be player 1, so  $v_1(\sigma) < v$ . Let  $(\sigma_1^*, \sigma_2^*)$  be a profile of minmax strategies. Note that  $v_1(\sigma_1^*, \sigma_2) \geq v$ . Since  $\sigma_2$  is not a minmax strategy, there exists a pure strategy  $s_1 \in \Delta_1$  such that  $v_1(s_1, \sigma_2) > v$ . Thus, for every  $\varepsilon \in (0, 2]$ , it holds that

$$v_1\left(\frac{\varepsilon}{2}s_1 + \left(1 - \frac{\varepsilon}{2}\right)\sigma_1^*, \sigma_2\right) > v.$$

It holds that

$$v_2\left(\frac{\varepsilon}{2}s_1 + \left(1 - \frac{\varepsilon}{2}\right)\sigma_1^*, \sigma_2^*\right) \geq -v,$$

so for every  $\varepsilon > 0$ ,  $f^2(\sigma)$  contains a state which is in an  $\varepsilon$ -neighborhood of a mixed strategy Nash equilibrium of  $G$ , and therefore  $f^\infty(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ .

- 1.2 Suppose  $v_1(\sigma) = v$ . Then, there exists a pure strategy  $s_1 \in \Delta_1$  such that

$$v_1(s_1, \sigma_2) > v,$$

since otherwise  $\sigma_2$  would be a minmax strategy. If  $s_1$  is a minmax strategy, then player 2 can deviate to a minmax strategy  $\sigma_2^*$  to obtain  $v_2(s_1, \sigma_2^*) = -v$ , i.e.,  $f^2(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ . If  $s_1$  is not a minmax strategy, then  $(s_1, \sigma_2) \in f_1(\sigma)$  is a state as in Case 1.1, so for every  $\varepsilon > 0$ ,  $f^3(\sigma)$  contains a state which is in a  $\varepsilon$ -neighborhood of a mixed strategy Nash equilibrium of  $G$ , and therefore  $f^\infty(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ .

**Case 2:**  $\sigma_1$  is a minmax strategy and  $\sigma_2$  is not, or  $\sigma_1$  is not a minmax strategy and  $\sigma_2$  is. Without loss of generality, assume  $\sigma_1$  is a minmax strategy.

- 2.1 If  $v_1(\sigma) > v$ , then player 2 can profitably switch to a minmax strategy  $\sigma_2^*$  and we are done.
- 2.2 If  $v_1(\sigma) = v$ , then since  $\sigma_2$  is not a minmax strategy, there exists a deviation to a pure strategy  $s_1 \in \Delta_1$  such that  $v_1(s_1, \sigma_2) > v$ . If  $s_1$  is a minmax strategy, then  $(s_1, \sigma_2) \in f_1(\sigma)$  is a state as in Case 2.1, so  $f_2(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ . If  $s_1$  is not a minmax strategy, then  $(s_1, \sigma_2) \in f_1(\sigma)$  is a state as in Case 1.1, and for every  $\varepsilon > 0$  it holds that  $f^3(\sigma)$  contains a state which is in an  $\varepsilon$ -neighborhood of a mixed strategy Nash equilibrium of  $G$ , so  $f^\infty(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ .

□

As a final result, we show the equivalence between the set of mixed strategy Nash equilibria of  $G$  and the MSS of the social environment  $\tilde{\Gamma}$  for finite two-player games where one of the two players has two pure strategies.

**Theorem 39.** Let  $\tilde{G}$  be the mixed-extension of a finite two-player game  $G$  and let  $\tilde{\Gamma}$  be the social environment corresponding to  $\tilde{G}$ . Assume that one player has two pure strategies in  $G$ . Then the MSS of  $\tilde{\Gamma}$  coincides with the set of mixed strategy Nash equilibria of  $G$ .



*Proof.* Assume without loss of generality that player 1 has two pure strategies. Let the set of pure strategies of player 1 be  $\{U, D\}$  with generic element  $A \in \{U, D\}$  and let the set of pure strategies of player 2 be given by  $\{s^1, \dots, s^\ell\}$  with generic element  $s^j$ . We also use the notation  $U$  and  $D$  for the mixed strategy that puts probability 1 on pure strategy  $U$  and  $D$ , respectively, and similarly for  $s^j$ .

Let some  $\sigma \in \Delta$  be given. By Corollary 37, it suffices to show the weak improvement property of  $\tilde{\Gamma}$ , i.e.,  $f^\infty(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ . We distinguish between two cases.

**Case 1:**  $G$  has a pure strategy Nash equilibrium, without loss of generality,  $(U, s^*)$ .

If  $\sigma$  is a mixed strategy Nash equilibrium of  $G$ , we are done, so assume  $\sigma$  is not a mixed strategy Nash equilibrium of  $G$ . If player 2 has a profitable deviation from  $\sigma$ , then there is a pure strategy best response  $s^j \in \Delta_2$  such that  $(\sigma_1, s^j) \in f(\sigma)$ . If  $(\sigma_1, s^j)$  is a mixed strategy Nash equilibrium of  $G$ , we are done. If not, then player 1 must have a pure strategy best response to  $(\sigma_1, s^j)$ , say  $A$ . Thus,  $f^2(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$  or a pure strategy profile  $(A, s^j)$ . The same conclusion holds if player 1 has a profitable deviation from  $\sigma$ . If the pure strategy profile  $(A, s^j)$  is a Nash equilibrium of  $G$ , we are done. If not, at least one player has a profitable deviation from it. We distinguish between two cases.

1.1  $A = D$ .

1.1.a Assume player 1 can profitably deviate from  $(D, s^j)$ . Then it holds that  $(U, s^j) \in f(D, s^j)$ . If  $(U, s^j)$  is a Nash equilibrium of  $G$ , we are done. If not, then player 2 can profitably deviate to the Nash equilibrium  $(U, s^*)$  of  $G$  and we are done.

1.1.b Assume player 2 can profitably deviate from  $(D, s^j)$ . Let  $s^h$  be a best response for player 2, so  $(D, s^h) \in f(D, s^j)$ . If this is a Nash equilibrium of  $G$ , we are done. Otherwise, player 1 can profitably deviate to  $(D, s^h)$ , which brings us back to Case 1.1.a.

1.2  $A = U$ .

1.2.a Assume player 2 can profitably deviate from  $(U, s^j)$ . It holds that the Nash equilibrium  $(U, s^*)$  of  $G$  belongs to  $f(U, s^j)$ , so we are done.

1.2.b Assume player 1 can profitably deviate from  $(U, s^j)$ . Then it holds that  $(D, s^j) \in f(U, s^j)$ . If  $(D, s^j)$  is a Nash equilibrium of  $G$ , then we are done. Else, player 2 must have a profitable deviation from  $(D, s^j)$ , which brings us back to Case 1.1.b.

**Case 2:**  $G$  has no pure strategy Nash equilibrium.

We first show that in every mixed strategy Nash equilibrium of  $G$ , player 1 plays both  $U$  and  $D$  with strictly positive probability. Towards a contradiction, suppose there is a mixed strategy Nash equilibrium  $(A, \sigma_2^*)$  of  $G$  such that player 1 plays a pure strategy, without loss of generality, strategy  $A = U$ . It holds that any pure strategy of player 2 in the support of  $\sigma_2^*$  is a best response against  $U$ . Since  $G$  has no pure strategy Nash equilibrium, it must hold that playing  $D$  against any pure strategy in the support of  $\sigma_2^*$  gives player 1 a strictly higher payoff than playing  $U$ . It follows that  $D$  is a profitable deviation for player 1 from  $(U, \sigma_2^*)$ . This contradicts  $(U, \sigma_2^*)$  being a mixed strategy Nash equilibrium of  $G$ .

To finish the proof, we show that  $f^\infty(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$ . As in the first part of Case 1, we can show that  $f^2(\sigma)$  contains a mixed strategy Nash equilibrium of  $G$  and we are done, or a pure strategy profile which is not a Nash equilibrium of  $G$ . Player 1 or player 2 has a profitable deviation from this pure strategy profile. In the latter case, player 2 can choose a pure strategy best response and in the next step, player 1 can profitably deviate to a pure strategy. In both cases it holds that there is  $k \in \mathbb{N}$  such that  $f^k(\sigma)$  contains a pure strategy profile  $(A, s^j)$  from which player 1 has a profitable deviation. Without loss of generality, let  $A = U$ .

Observe that for player 1 any completely mixed strategy is a profitable deviation from  $(U, s^j)$ . Let  $\sigma^*$  be a mixed strategy Nash equilibrium of  $G$  and let  $p \in (0, 1)$  denote the probability that  $\sigma_1^*$  puts on  $U$ . We distinguish 3 cases.

$$2.1 \quad v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*) > v_2(D, s^j) - v_2(U, s^j).$$

For  $\varepsilon \in (0, p)$ , let  $\sigma'_1$  be the strategy where player 1 plays  $U$  with probability  $p - \varepsilon/2$ . Since any completely mixed strategy of player 1 is a profitable deviation from  $(U, s^j)$ , it holds that  $(\sigma'_1, s^j) \in f(U, s^j)$ . We have that

$$\begin{aligned} v_2(\sigma'_1, s^j) &= v_2(\sigma_1^*, s^j) + \frac{\varepsilon}{2}(v_2(D, s^j) - v_2(U, s^j)) \\ &< v_2(\sigma^*) + \frac{\varepsilon}{2}(v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*)) \\ &= v_2(\sigma'_1, \sigma_2^*), \end{aligned}$$

where the strict inequality uses that  $\sigma_2^*$  is a best response against  $\sigma_1^*$  and the assumption of Case 2.1. It follows that  $(\sigma'_1, \sigma_2^*) \in f(\sigma'_1, s^j)$ . Since  $\varepsilon > 0$  can be chosen arbitrarily small, this shows that  $\sigma^* \in f^\infty(\sigma)$ .

$$2.2 \quad v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*) < v_2(D, s^j) - v_2(U, s^j).$$

For  $\varepsilon \in (0, 1-p)$ , let  $\sigma'_1$  be the strategy where player 1 plays  $U$  with probability  $p + \varepsilon/2$ . The proof now follows as in Case 2.1.

$$2.3 \quad v_2(D, \sigma_2^*) - v_2(U, \sigma_2^*) = v_2(D, s^j) - v_2(U, s^j).$$

It holds that  $(D, s^j) \in f(U, s^j)$ .

Let  $s^h$  be a best response of player 2 against  $D$  and, for  $\varepsilon \in (0, 1)$ , let  $\sigma'_2$  be the strategy that puts weight  $(1 - \varepsilon)$  on  $\sigma_2^*$  and weight  $\varepsilon$  on  $s^h$ . We have that

$$\begin{aligned} v_2(D, \sigma'_2) &= v_2(\sigma^*) + pv_2(D, \sigma_2^*) - pv_2(U, \sigma_2^*) \\ &\geq v_2(\sigma_1^*, s^j) + pv_2(D, s^j) - pv_2(U, s^j) \\ &= v_2(D, s^j), \end{aligned} \tag{A.1}$$

where the inequality uses that  $\sigma^*$  is a mixed strategy Nash equilibrium of  $G$  and the assumption of Case 2.3. Since  $(D, s^j)$  is not a Nash equilibrium of  $G$ , it holds that  $v_2(D, s^h) > v_2(D, s^j)$ . By (A.1) and the definition of  $\sigma'_2$ , it now follows that  $v_2(D, \sigma'_2) > v_2(D, s^j)$ , so  $(D, \sigma'_2) \in f(D, s^j)$ . Since  $(D, s^h)$  is not a Nash equilibrium of  $G$  and  $s^h$  is a best response against  $D$ , we have that  $v_1(\sigma_1^*, s^h) > v_1(D, s^h)$ . It follows that

$$v_1(\sigma_1^*, \sigma'_2) = (1 - \varepsilon)v_1(\sigma^*) + \varepsilon v_1(\sigma_1^*, s^h) > (1 - \varepsilon)v_1(D, \sigma_2^*) + \varepsilon v_1(D, s^h) = v_1(D, \sigma'_2),$$

so  $(\sigma_1^*, \sigma'_2) \in f(D, \sigma'_2)$ . Since  $\varepsilon > 0$  can be chosen arbitrarily small, we have that  $\sigma^* \in f^\infty(\sigma)$ , which concludes the proof.

□

We analyzed the game of matching pennies in Example 4 and concluded that better-response dynamics did not single out any strategy profile. The game of matching pennies satisfies the assumptions of both Theorems 38 and 39. The MSS of this game therefore consists of the unique mixed strategy Nash equilibrium where each pure strategy is played with probability  $1/2$ .

# Chapter 3



# Bertrand Competition with Asymmetric Costs: A Solution in Pure Strategies

## 3.1 Introduction

The analysis of price competition is a fundamental part of oligopoly theory since Bertrand's contribution (Bertrand, 1883). The Bertrand duopoly with symmetric constant marginal costs, homogeneous goods and continuous prices has a unique pure strategy Nash Equilibrium characterized by a strategy profile in which prices equal marginal costs.

If marginal costs are not symmetric across firms and the market is shared if firms set equal prices, no pure strategy Nash equilibrium exists. Blume (2003) shows that there exists a Nash equilibrium in mixed strategies where the more efficient firm sets price equal to the opponent's marginal cost and serves the entire market with probability 1. The rival randomizes on an interval above his marginal cost. Kartik (2011) strengthens the result by showing that all undominated equilibria lead to the same market price and shares. The complete set of undominated Nash equilibria is not constructed in these chapters.

In this chapter, we analyze the same Bertrand game. We follow Blume (2003) and Kartik (2011) in that we focus on strategy profiles in which no firm chooses a predatory price, i.e., a price below marginal cost. Instead of Nash equilibrium, however, we employ a solution concept recently introduced by Demuynck, Herings, Saulle, and Seel (2019), the Myopic Stable Set. A set of strategy profiles is myopically stable if it satisfies three conditions, *deterrence of external deviations*, *asymptotic external stability* and *minimality*. Deterrence of external deviations requires that no player benefits by switching her strategy such that the resulting strategy profile is outside the Myopic Stable Set. Asymptotic external stability makes sure that from any strategy profile outside the set it is possible to get arbitrarily close to a strategy profile inside the Myopic Stable Set by a sequence of better replies. Minimality requires that the Myopic Stable Set is minimal with respect to set inclusion.

In Demuynck, Herings, Saulle, and Seel (2019) we defined the Myopic Stable Set for a very general class of *social environments* (Chwe, 1994) that allows for infinite state spaces and includes normal-form games as a special case. We proved that if the state space is compact then the Myopic Stable Set exists and under some mild continuity assumptions it is also unique. Moreover, we showed that the Myopic Stable Set coincides with the set of pure strategy Nash Equilibria for supermodular games, aggregative games and potential games.

In light of these results, the Bertrand model with asymmetric costs is interesting for several reasons: it does not satisfy the compactness and continuity assumptions of De-

muynck, Herings, Saulle, and Seel (2019), it does not belong to any of the aforementioned classes of games and the set of pure strategy Nash equilibria of this game is empty. Moreover, the set of mixed strategy Nash equilibria is large.

We prove existence and uniqueness of the Myopic Stable Set for symmetric and asymmetric Bertrand competition. We characterize the set in closed form. The set is small and gives an intuitive prediction.

## 3.2 Model and Solution Concept

In this section, we define the Myopic Stable Set for normal-form games  $G := (N, (S_i)_{i \in N}, (\pi_i)_{i \in N})$ , where  $N = \{1, \dots, n\}$  is a finite set of players with typical element  $i$ ,  $S_i$  is the set of pure strategies for each player  $i \in N$  and  $\pi_i : \times_{i \in N} S_i \rightarrow \mathbb{R}$  is the payoff function of player  $i$ . Denote the strategy space by  $S = \times_{i \in N} S_i \subseteq \mathbb{R}^n$ . We use the standard notation  $s := (s_i, s_{-i}) \in S$  to denote the strategy profile where  $s_{-i}$  is the list of strategies of all players except  $i$ , i.e.,  $s_{-i} := (s_j)_{j \in N \setminus \{i\}}$ .

We say that a strategy profile  $s' \in S$  dominates another strategy profile  $s \in S$  if there is a player who can unilaterally deviate to  $s'$  and strictly prefers  $s'$  over  $s$ , i.e., from  $s$ , the player has a better reply such that the resulting strategy profile is  $s'$ .

**Definition 40 (Dominance).** Let  $s, s' \in S$  be two strategy profiles. The strategy profile  $s'$  dominates  $s$  if there exists a player  $i \in N$  such that  $\pi_i(s') > \pi_i(s)$  and  $s'_{-i} = s_{-i}$ .

Let some strategy profile  $s \in S$  be given. The set of all strategy profiles that dominates  $s$  together with  $s$  itself is denoted by  $f(s)$ ,

$$f(s) = \{s\} \cup \{s' \in S \mid s' \text{ dominates } s\}.$$

We define the two-fold composition of  $f$  by

$$f^2(s) = \{s'' \in S \mid \exists s' \in S : s' \in f(s) \text{ and } s'' \in f(s')\}.$$

By induction, for  $k \geq 3$ , we can define the  $k$ -fold composition  $f^k(s)$  by  $s^k \in f^k(s)$  if there is  $s^{k-1} \in f^{k-1}(s)$  such that  $s^k \in f(s^{k-1})$ . Observe that for all  $k, \ell \in \mathbb{N}$  such that  $k \leq \ell$  it holds that  $f^k(s) \subseteq f^\ell(s)$ . We define the set of all strategy profiles that can be reached from  $s$  by a finite number of dominations by  $f^{\mathbb{N}}(s)$ , where

$$f^{\mathbb{N}}(s) := \bigcup_{k \in \mathbb{N}} f^k(s).$$

Given  $s, s' \in S$ , we say that a strategy profile  $s'$  asymptotically dominates  $s$  if starting from  $s$ , it is possible to get arbitrarily close to  $s'$  in a finite number of dominations. Let  $d(s, s')$  denote the Euclidean distance between  $s$  and  $s'$ . The asymptotic dominance criterion is formally defined as follows:

**Definition 41 (Asymptotic Dominance).** A strategy profile  $s^\infty \in S$  asymptotically dominates  $s \in S$  if for all  $\varepsilon > 0$  there exists  $s' \in f^{\mathbb{N}}(s)$  such that  $d(s^\infty, s') < \varepsilon$ .

We denote the set of all strategy profiles in  $S$  that asymptotically dominate  $s$  by  $f^\infty(s)$ . Formally,

$$f^\infty(s) = \{s^\infty \in S \mid \forall \varepsilon > 0, \exists s' \in f^{\mathbb{N}}(s) : d(s^\infty, s') < \varepsilon\}.$$

It is easy to see that the set  $f^\infty(s)$  coincides with the closure of the set  $f^{\mathbb{N}}(s)$ . We are now ready to define the Myopic Stable Set, abbreviated as MSS, for normal-form games.

**Definition 42** (Myopic Stable Set (MSS)). Let  $G := (N, (S_i)_{i \in N}, (\pi_i)_{i \in N})$  be a normal-form game. The set  $M \subseteq S$  is a *Myopic Stable Set (MSS)* if it is closed and satisfies the following three conditions:

1. **Deterrence of external deviations:** For all  $s \in M$ ,  $f(s) \subseteq M$ .
2. **Asymptotic external stability:** For all  $s \notin M$ ,  $f^\infty(s) \cap M \neq \emptyset$ .
3. **Minimality:** There is no closed set  $M' \subsetneq M$  that satisfies Conditions 1 and 2.

Let  $M$  be a MSS. Deterrence of external deviations requires that no player can profitably deviate to a strategy profile outside  $M$ . Asymptotic external stability requires that any strategy profile outside  $M$  is asymptotically dominated by a strategy profile in  $M$ . Hence, from any strategy profile outside of  $M$  it is possible to get arbitrary close to a strategy profile in  $M$  by a finite number of dominations. Observe that an empty set would necessarily violate asymptotic external stability, so it follows that  $M$  is non-empty.

While we focus on pure strategies, the MSS can be defined analogously on the set of mixed strategies; for details, see Online Supplement A.6 of Demuynck, Herings, Saulle, and Seel (2019). In any (mixed) Nash equilibrium, we have  $f(s) = f^\infty(s) = s$ . Thus, by asymptotic external stability, all mixed-strategy Nash equilibria are part of the MSS in mixed strategies.

### 3.3 Bertrand Duopoly

Consider a model with two firms,  $N = \{1, 2\}$ , having constant marginal costs  $0 \leq c_1 \leq c_2$  and no fixed costs. In many countries, pricing below marginal or average cost is considered to be predatory pricing and is thus forbidden by law. Thus, we require the strategy space  $P_i$  of a firm  $i \in N$  to consist of all non-predatory prices, i.e.,  $P_i = [c_i, \infty)$  and we define  $P := P_1 \times P_2$  with typical element  $p$ .

The absence of predatory pricing is similar to the focus on undominated equilibrium strategies in Kartik (2011). However, pricing at marginal costs is allowed in our setting, while it is ruled out in equilibria considered by Kartik (2011). This difference does not influence the set of undominated Nash equilibria in the asymmetric case  $c_1 < c_2$ , as, by a standard argument in this literature, no equilibrium in which Firm 2 places mass at  $c_2$  can exist. However, restricting attention to equilibria with price strictly larger than marginal cost leads to non-existence in the symmetric case which is not particularly appealing.

The continuous market demand function  $D : [0, \infty) \rightarrow \mathbb{R}_+$  expresses the demand for the good as a function of the lowest price  $p_m$  available in the market. There exists a ‘‘choke’’ price  $\bar{p}_m \in (c_2, \infty)$  such that  $D(p_m) = 0$  for all  $p_m \geq \bar{p}_m$ . The demand is strictly decreasing



on  $[0, \bar{p}_m]$  and zero on  $[\bar{p}_m, \infty)$ . We assume that there is an  $\varepsilon > 0$  such that the profits  $(p_m - c_1)D(p_m)$  of a monopolist with marginal costs  $c_1$  are strictly increasing in  $p_m$  on the domain  $[c_1, c_2 + \varepsilon]$ . This assumption ensures an element of competition between the firms. If the profit of a monopolist with marginal cost  $c_1$  has a maximum for  $p_1 \in [c_1, c_2)$ , then the strategy profiles in which Firm 1 chooses  $p_1$  and Firm 2 chooses a strictly higher price are Nash equilibria and part of the MSS.

For each firm  $i \in \{1, 2\}$ , its individual demand depends on its price  $p_i$  and the price of its competitor  $j \neq i$  and is given by:

$$Q_i(p_i, p_j) = \begin{cases} D(p_i) & \text{if } p_i < p_j, \\ D(p_i)/2 & \text{if } p_i = p_j, \\ 0 & \text{if } p_i > p_j. \end{cases}$$

Thus, the firm setting the lower price serves the entire market. In case of a tie, the market share of each firm equals  $1/2$ . For each firm  $i \in \{1, 2\}$ , the demand  $Q_i(p_i, p_j)$  leads to a payoff function  $\pi_i : P_1 \times P_2 \rightarrow \mathbb{R}$  defined by  $\pi_i(p_i, p_j) = (p_i - c_i)Q_i(p_i, p_j)$ . The normal-form game  $(\{1, 2\}, (P_i)_{i \in \{1, 2\}}, (\pi_i)_{i \in \{1, 2\}})$  is denoted by  $\Gamma$ .

For  $c_1 < c_2$ , let  $\underline{p}_1 \in (c_1, c_2)$  be the price such that  $\pi_1(\underline{p}_1, c_2) = \pi_1(c_2, c_2)$ . To see that this price is uniquely determined, note that  $\pi_1(c_1, c_2) = 0$ ,  $\pi_1(c_2, c_2) > 0$ , and  $\pi_1(p_1, c_2)$  is strictly increasing and continuous for  $p_1 \in [c_1, c_2)$  and has a downward jump at  $c_2$ . By the intermediate value theorem,  $\underline{p}_1$  exists. Since  $\pi_1(p_1, c_2)$  is strictly increasing for  $p_1 \in [c_1, c_2)$ , we obtain that  $\underline{p}_1$  is uniquely determined. This allows us to characterize the MSS as follows:

**Proposition 43.** (i) If  $c_1 = c_2 = c$ , the unique MSS of  $\Gamma$  is given by  $P^* = \{p_1 = p_2 = c\}$ .

(ii) If  $c_1 < c_2$ , the unique MSS of  $\Gamma$  is given by

$$P^* = \{(p_1, p_2) \in P \mid \pi_1(p_1, c_2) \geq \pi_1(c_2, c_2), p_2 = c_2\} = [\underline{p}_1, c_2] \times \{c_2\}.$$

*Proof.* We first prove Part (i). Deterrence of external deviations is satisfied as  $\pi_1(c, c) = \pi_2(c, c) = 0$  and  $\pi_1(p_1, c) \leq 0, \pi_2(c, p_2) \leq 0$ . As  $f^\infty(c, c) = \{(c, c)\}$ , the profile  $(c, c)$  is part of every MSS. Hence, by minimality, if  $P^*$  is a MSS, then it is also unique.

It remains to verify asymptotic external stability. Without loss of generality, let  $p_1 \leq p_2$ . Let some  $(p_1, p_2) \neq (c, c) \in P$  be given. If  $c < p_1 < p_2$ , then Firm 2 can profitably deviate to  $p'_2 = (c + \min\{p_1, \bar{p}_m\})/2$  and Firm 1 can profitably deviate in the next step to  $p'_1 = (c + p'_2)/2$  and so forth. It follows that  $(c, c) \in f^\infty(p_1, p_2) \cap P^*$ . If  $p_1 = c < p_2$ , then Firm 1 can profitably deviate to  $(p'_1, p_2)$ , where  $p'_1 \in (c, p_2)$  is chosen such that  $p'_1 < c + \varepsilon$  with  $\varepsilon > 0$  such that the profits  $(p_m - c)D(p_m)$  of a monopolist with marginal costs  $c$  are strictly increasing in  $p_m$  on the domain  $[c, c + \varepsilon]$ . We can then continue as in the previous case. If  $p_1 = p_2 > c$ , then there is  $p'_1 \in (c, p_2)$  such that  $\pi_1(p'_1, p_2) > \pi_1(p_1, p_2)$ . Either  $p_2 \leq \bar{p}_m$  and  $p'_1$  is obtained by slightly undercutting  $p_2$ , or  $p_2 > \bar{p}_m$  and any  $p'_1 \in (c, \bar{p}_m)$  will do. Now we can continue as before. Because of the legal restrictions on predatory pricing, we have covered all strategy profiles in  $P \setminus P^*$  and thereby shown that  $P^*$  satisfies asymptotic external stability which completes the proof of (i).

We split the proof of (ii) into several steps. The first step verifies deterrence of external deviations and asymptotic external stability. The remaining steps establish minimality and uniqueness.

Step 1. We show that  $P^*$  satisfies deterrence of external deviations and asymptotic external stability. We start with deterrence of external deviations. For  $(p_1, c_2) \in P^*$ , profits of Firm 1 are greater than or equal to  $\pi_1(c_2, c_2)$ . For  $p_1 \notin [\underline{p}_1, c_2]$ , profits are  $\pi_1(p_1, c_2) < \pi_1(c_2, c_2)$ . Thus, for any state in  $P^*$ , player 1 will not deviate to a state in  $P \setminus P^*$ . Firm 2 makes zero profits for any point in  $P^*$  and non-positive profits for any  $p_2 \in P_2$  if  $p_1 \leq c_2$ . Thus, Firm 2 has no profitable deviation. This shows deterrence of external deviations for  $P^*$ .

It remains to verify asymptotic external stability. Let some  $(p_1, p_2) \in P \setminus P^*$  be given. If  $c_2 < p_1 < p_2$ , then Firm 2 can profitably deviate to  $p'_2 = (c_2 + \min\{p_1, p_m\})/2$  and Firm 1 can profitably deviate in the next step to  $p'_1 = (c_2 + p'_2)/2$  and so forth. It follows that  $(c_2, c_2) \in f^\infty(p_1, p_2) \cap P^*$ . If  $p_1 \leq c_2 < p_2$ , then Firm 1 can profitably deviate to  $(p'_1, p_2)$ , where  $p'_1 \in (c_2, p_2)$  is chosen such that  $p'_1 < c_2 + \varepsilon$  with  $\varepsilon > 0$  such that the profits  $(p_m - c_1)D(p_m)$  of a monopolist with marginal costs  $c_1$  are strictly increasing in  $p_m$  on the domain  $[c_1, c_2 + \varepsilon]$ . We can then continue as in the previous case. If  $c_2 < p_2 \leq p_1$ , then there is  $p'_1 \in (c_2, p_2)$  such that  $\pi_1(p'_1, p_2) > \pi_1(p_1, p_2)$ . Either  $p_2 \leq \bar{p}_m$  and  $p'_1$  is obtained by slightly undercutting  $p_2$ , or  $p_2 > \bar{p}_m$  and any  $p'_1 \in (c_2, \bar{p}_m)$  will do. Now we can continue as before. If  $p_1 \notin [\underline{p}_1, c_2]$  and  $p_2 = c_2$ , then Firm 1 can profitably deviate to  $c_2$  to reach a strategy profile in  $P^*$ . Because of the legal restrictions on predatory pricing, we have covered all strategy profiles in  $P \setminus P^*$  and thereby shown that  $P^*$  satisfies asymptotic external stability.

Step 2. Let  $M \subseteq P^*$  be a set satisfying deterrence of external deviations and asymptotic external stability. Let us show that for every  $p_1 \in [\underline{p}_1, c_2)$  it holds that  $(p_1, c_2) \in M$  implies  $(c_2, c_2) \in M$ . Suppose  $(p_1, c_2) \in M$  and  $(c_2, c_2) \notin M$ . By closedness of  $M$ , there is  $\varepsilon' > 0$  such that, for every  $\varepsilon \in (0, \varepsilon')$ ,  $(c_2 - \varepsilon, c_2) \notin M$ . Take  $p'_1 = \max\{(p_1 + c_2)/2, c_2 - \varepsilon'/2\}$ , then  $(p'_1, c_2) \in f(p_1, c_2)$ , so  $(p'_1, c_2) \in M$ . Given that  $p'_1 > c_2 - \varepsilon'$ , we obtain a contradiction.

Step 3. Let  $M \subseteq P^*$  be a set satisfying deterrence of external deviations and asymptotic external stability. Let us show that if  $(c_2, c_2) \in M$ , then, for every  $p_1 \in [\underline{p}_1, c_2)$ , we have  $(p_1, c_2) \in M$ . This follows from the fact that, for every  $p_1 \in (\underline{p}_1, c_2)$ ,  $\pi_1(p_1, c_2) > \pi_1(c_2, c_2)$  and the fact that  $M$  is closed.

Step 4. We are now ready to show that  $P^*$  is an MSS. First of all, by Step 1 it satisfies deterrence of external deviations and asymptotic external stability. So if  $P^*$  is not an MSS, it should violate minimality. This means that there is a proper subset  $M$  of  $P^*$  that also satisfies deterrence of external deviations and asymptotic external stability. The set  $M$  either contains  $(c_2, c_2)$  or it is a subset of  $[\underline{p}_1, c_2) \times \{c_2\}$ . If  $M$  contains  $(c_2, c_2)$  then, by Step 3, it should contain  $[\underline{p}_1, c_2) \times \{c_2\}$  and therefore be equal to  $P^*$ , a contradiction with  $M$  being a proper subset of  $P^*$ . If  $M$  is a subset of  $P^* \setminus \{(c_2, c_2)\}$ , then by Step 2, it should contain  $(c_2, c_2)$ , a contradiction.

Step 5. Finally we show that  $P^*$  is the unique MSS. Let  $M$  be an MSS. We show that  $P^* \cap M \neq \emptyset$ . Towards a contradiction, suppose that  $P^* \cap M = \emptyset$ . Then, for all  $p' \in P^*$ , there is  $p'' \in M$  such that  $p'' \in f^\infty(p')$ . Given that  $P^*$  is closed and  $p'' \notin P^*$ , there is  $\varepsilon > 0$  such that  $B_\varepsilon(p'') \cap P^* = \emptyset$ , where  $B_\varepsilon(p'') = \{p \in P \mid d(p, p'') < \varepsilon\}$ . By definition of  $f^\infty$ , there is  $k \in \mathbb{N}$  and  $p \in P$  such that  $p \in f^k(p')$  and  $p \in B_\varepsilon(p'')$ . By a  $k$ -fold application

of deterrence of external deviations, it holds that  $p \in P^*$ , so  $p \in B_\varepsilon(p'') \cap P^*$  and we have obtained a contradiction. Consequently, it holds that  $P^* \cap M \neq \emptyset$ .

Step 6. If  $M$  contains  $(c_2, c_2)$ , then, by Step 3,  $M$  should also contain  $P^* \setminus \{(c_2, c_2)\}$ , so  $P^* \subseteq M$  and by minimality  $P^* = M$ . If  $M$  contains an element of  $P^* \setminus \{(c_2, c_2)\}$ , then, by Step 2, it should also contain  $(c_2, c_2)$  and, by Step 3, also  $P^* \setminus \{(c_2, c_2)\}$ . Again, we obtain  $P^* \subseteq M$  and by minimality  $P^* = M$ .

□

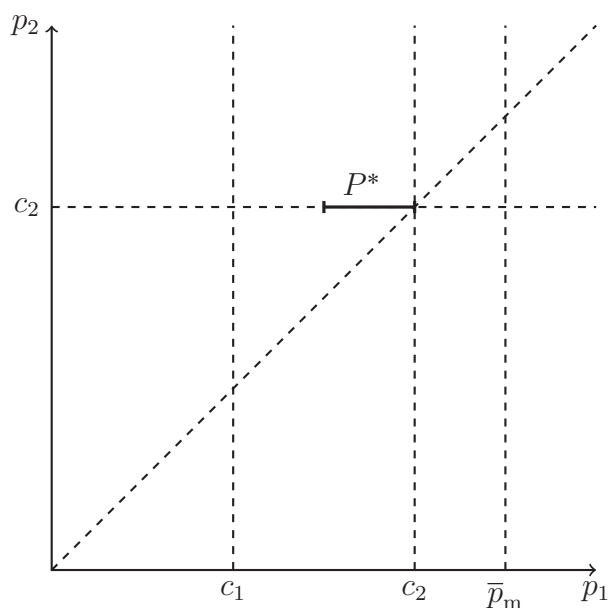
For the case  $c_1 = c_2 = c$ , the MSS yields a unique prediction which coincides with the well-known reasoning that prices are equal to marginal cost. This is also the pure strategy Nash equilibrium of the game.

For  $c_1 < c_2$ , the Myopic Stable Set is illustrated in Figure 3.1. The set is relatively small. In any element of the set, firm 2 chooses price equal to its marginal cost. Firm 1 can choose any price between  $\underline{p}_1$  and  $c_2$ . To provide an intuition, we start with the point  $(c_2, c_2)$ . From this point,  $p_1 \in (\underline{p}_1, c_2)$  are all better responses (with  $\underline{p}_1$  included by closedness). From these points, increasing  $p_1$  to a value below  $c_2$  are again all better responses. Finally,  $(c_2, c_2)$  is included by closedness. The proof shows that from any point outside the set, there is a sequence of myopic improvements that approaches the set.

In contrast to the unique MSS, the complete set of Nash equilibria has not been characterized. However, as Kartik (2011) shows, there are infinitely many Nash equilibria in mixed strategies which share the property that Firm 1 chooses  $p_1 = c_2$  and Firm 2 chooses a mixed strategy with support starting at  $c_2$ . Thus, for the Bertrand game without predatory pricing, the MSS provides intuitive predictions which are less ambiguous than the Nash equilibrium predictions.

In all profiles contained in the MSS, the profit of Firm 1 is strictly lower than its Nash equilibrium profit. Among the profiles in the MSS, the profit of Firm 1 increases as  $p_1$  approaches  $c_2$  from below. In the limit, the profit of Firm 1 even approaches its Nash equilibrium profit.

Figure 3.1: The MSS for the asymmetric Bertrand model.



### 3.4 Discussion

We have provided a complete characterization of the Myopic Stable Set for the Bertrand duopoly with asymmetric marginal costs. Despite the popularity of Bertrand competition, the set of Nash equilibria has not been fully characterized, but infinitely many equilibria exist.

The Myopic Stable Set yields a unique set-valued prediction. The market price is predicted to be weakly lower than in any Nash equilibria in undominated strategies. A clean solution to a game with such a complex set of Nash equilibria boosts the appeal of the Myopic Stable Set as a solution concept.



# Chapter 4



# The Last Will Be the First and The First Last: Segregation in Societies with Positional Externalities

## 4.1 Introduction

Coalition formation is an important pattern of human behavior and many social phenomena can be studied in terms of a coalition formation process. For example, think of the provision of a local public good, the bargaining process underlying environmental agreements, political party formation, cartel formation, and racial integration. Typically, we think about the coalition formation process as a voluntary one. And indeed, in coalitional games, we mostly assume that individuals cannot exert their power or abilities in order to subjugate others.<sup>1</sup>

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

In the state of nature, however, individuals share the wish of ending the conflicts in order to claim properties and to benefit from their goods. This common desire leads into a “social contract” which bans the coercive force in favor of the enforcement power of institutions which ensure social norms and property rights.<sup>2</sup> In broad economic terms, different societies can be seen as the outcome of different social contracts.

In this chapter, we model two possible and alternative outcomes of a social contract, a competitive and an egalitarian society. The competitive society is characterized by the fact that it does not allow coercion, but there is no enforcement of resource redistribution within it, i.e., each member of a coalition receives a material payoff which reflects the member’s ability. The egalitarian society also does not allow for coercive force, but guarantees through institutions or norms an equal distribution of resources within a coalition, i.e., each member of a coalition receives the same material payoff.

---

<sup>1</sup>Notable exceptions of models with different forms of coercion are due to Piccione and Rubinstein (2007), Acemoglu, Egorov and Sonin (2008), and Piccione and Razin (2009).

<sup>2</sup>On this point of view, see Muthoo (2004) and Hafer (2006) who study models explaining the rise of property rights from the state of nature.



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

Following the same line of reasoning, Frank (1985) argues that positional concerns can explain many real world phenomena such as flatter intrafirm wage profiles in which low productivity workers and high productivity workers are paid more and less with respect to their marginal contribution to compensate for the difference in relative ranking. More recently, Card, Mas, Moretti, and Saez (2012) have shown that high relative wages affect job satisfaction and Bracha, Gneezy, and Loewenstein (2015) have shown a positive correlation between the supply of labor and relative wages. Moreover, happiness is found to be significantly and negatively affected by a lower relative income (Ferrer-i-Carbonell, 2005; Luttmer, 2005; Clark, Frijters, and Shields, 2008). To sum up, it is widely accepted that people do not only care about their absolute wealth, but also about their relative ranking in society.

Motivated by these findings, we propose a coalition formation model with positional externalities. A state in this model corresponds to a particular partition of the individuals into coalitions. We model the importance of the relative position in a coalition by the intragroup relative payoff. This measures the difference between the material payoff of an individual and the average material payoff of the coalition to which the individual belongs.

For the analysis of competitive and egalitarian societies, we introduce two notions of segregation, *bottom-up segregation* and *top-down segregation*. Roughly speaking, a state is bottom-up segregated if every coalition is formed by individuals that are adjacent in abilities and only the most productive individuals may not be part of any productive coalition. On the opposite, a state is top-down segregated if every coalition is formed by individuals with adjacent productivities, but only the weakest individuals may not be part of any productive coalition.

We employ two different solution concepts to analyze the different societies. First, the classic concept of the Core which consists of all allocations for which there is no coalition such that all members benefit from deviating. As a second solution concept, we consider the Myopic Stable Set as introduced in Demuynck, Herings, Saulle, and Seel (2018). In the setting of this chapter, the Myopic Stable Set is the minimal set of states such that (i) from any coalition structure in the set, no sequence of myopic coalitional improvements leaves the set and (ii) for any given initial coalition structure, there is a sequence of myopic coalitional improvements which leads to an element in the set.

For both competitive and egalitarian societies, we provide a full characterization of the Core and the Myopic Stable Set and show that both solution concepts coincide. In particular, we prove that for competitive societies the Myopic Stable Set consists of the set of bottom-up segregated states and for egalitarian societies the Myopic Stable Set consists of the set of top-down segregated states. This finding might seem counterintuitive at first glance as one might expect that the most advantaged individuals are best-off in competi-

tive societies and the least advantaged individuals are best-off in egalitarian societies.

The reasoning behind the predictions for competitive societies is as follows. First, note that the individuals with the lowest productivity are very attractive as coalition members, since they yield a higher relative payoff to other members of the coalition. Since these incentives are mutual, groups of individuals with low productivity will form. Intuitively, a similar process repeats until only individuals with high productivity might remain in unproductive coalitions. By contrast, the results for egalitarian societies rely on the fact that in an egalitarian society there is no envy among individuals, i.e. the relative payoff of each individual in each coalition is the same. Thus, individuals prefer to stay with highly productive individuals in order to increase their material payoff.

We conclude by providing a comparative welfare analysis between the two types of societies. When all individuals differ in productivity, stable states in egalitarian societies are always at least as materially efficient as stable states in competitive societies. This is caused by the fact that in an egalitarian society, the most productive individuals are always part of a productive coalition, while this might not be the case for competitive societies. We show that in the general case, with arbitrary productivities, there can be stable states in a competitive society which lead to strictly higher material efficiency than all stable states in egalitarian societies.

### 4.1.1 Related Literature

This chapter is part of the vast literature on coalition formation theory (e.g. Ray, 2007; Ray and Vohra, 2014) and it is in line with contributions on hedonic games (e.g. Banerjee, Konishi, and Sönmez, 2001; Bogomolnaia and Jackson, 2002), social status (Milchtaich and Winter, 2002), and segregation (e.g. Goyal, Hernández, Martínez-Cánovas, Moisan, Muñoz-Herrera, and Sánchez, 2018).

In particular, our work is closely related to an increasing branch of literature devoted to the role of social norms in group formation (e.g. Watts, 2007; Piccione and Razin, 2009; Barberà, Beviá, and Ponsatí, 2015; Morelli and Park, 2016; Beviá, Córchon, and Romero-Medina, 2017). As in Piccione and Razin (2009), Morelli and Park (2016), and Barberà, Beviá, and Ponsatí (2015), we consider a model of coalition formation in which agents are vertically differentiated in productivity.

Piccione and Razin (2009) study a partition function game in which the relative strength of coalitions is represented by an exogenous “power relation”. Agents care only about their social ranking, which is determined by their own individual power and by the power of the coalition to which they belong.

Morelli and Park (2016) study how the degree of heterogeneity in the remuneration of the individuals affects the number of rival coalitions. The heterogeneity in remuneration is modeled by a ranking rule defined on the imputation rule. Moreover, they distinguish between productivity and material efficiency defining the last one as an increasing function with respect to the cardinality of the coalition. In contrast to Morelli and Park (2016), we define the ranking rule on productivity levels.

In line with Barberà, Beviá, and Ponsatí (2015), we consider a stylized model of a society consisting of individuals, a productivity vector, and a productivity threshold. We study two possible allocation rules, the egalitarian one and the competitive one. In contrast to Barberà, Beviá, and Ponsatí (2015), we incorporate positional concerns into our analysis.

Since we implement a ranking rule in the preference structure, our contribution can be also broadly related to the literature on social preferences (e.g. Watts, 2007; Dufwenberg, Heidhues, Kirchsteiger, Riedel, and Sobel, 2011; Maccheroni, Marinacci, and Rustichini, 2012). In particular, from this point of view, it is in line with Watts (2007), who studies two different settings. In one of them, individuals prefer coalitions with high-skilled members and in the other individuals prefer coalitions with low-skilled members. In Watts (2007), the number of coalitions is fixed and ties are broken in favor of the biggest coalition.

## 4.2 Model and Key Definitions

We first describe the building blocks of a *society*  $(N, \lambda, \rho)$ . The finite set  $N = \{1, \dots, n\}$  consists of all individuals in the society. A coalition  $S$  is a non-empty subset of  $N$  and the collection of all possible non-empty coalitions is denoted by  $\mathcal{N} = 2^N \setminus \{\emptyset\}$ . The collection of all partitions of a coalition  $S \in \mathcal{N}$  is denoted by  $\mathbb{P}(S)$ . The set  $\mathbb{P}(N)$  is denoted by  $X$  and is called the state space. An element  $x \in X$  is called a state and we write  $|x|$  for the number of coalitions in  $x$ .

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

The *productivity threshold*  $\rho \in \mathbb{N}$  denotes the minimal level below which a coalition is not productive. To make the problem interesting, we assume throughout the chapter that  $\rho \leq n$ . The *surplus*  $u(S)$  of coalition  $S \in \mathcal{N}$  equals  $u(S) = \sum_{i \in S} \lambda_i$  if  $|S| \geq \rho$  and 0 otherwise. The *average surplus* in coalition  $S$  is denoted by  $\bar{u}(S) = u(S)/|S|$ . To get an intuition for the surplus function, think of the coalition as a firm which produces using a production technology which requires a minimal number of workers to produce any output and exhibits constant returns to scale above this threshold.<sup>3</sup>

Our next step is to specify how a coalition  $S \in \mathcal{N}$  distributes its surplus over the coalition members. We denote the *material payoff* of individual  $i$  in coalition  $S$  by  $m_i(S) \in \mathbb{R}$ . Members of unproductive coalitions have a material payoff equal to zero. For members of productive coalitions, we consider two natural ways to divide the surplus. A division is *competitive* if each member of a productive coalition is paid according to productivity. A division is *egalitarian* if each member of a productive coalition receives the same payoff. These restrictions could be legal requirements, but they can also be thought of as social norms within a society.

**Definition 44** (Competitive Society). A society  $(N, \lambda, \rho)$  is *competitive* if surplus division within coalitions is competitive, i.e., for all  $S \in \mathcal{N}$  and for all  $i \in S$ ,  $m_i(S) = \lambda_i$  if  $|S| \geq \rho$  and  $m_i(S) = 0$  otherwise.

<sup>3</sup>An alternative specification would be to define the threshold in terms of a minimum sum of productivities. This would lead to similar results.

**Definition 45** (Egalitarian Society). A society  $(N, \lambda, \rho)$  is *egalitarian* if surplus division within coalitions is egalitarian, i.e., for all  $S \in \mathcal{N}$  and for all  $i \in S$ ,  $m_i(S) = \bar{u}(S)$  if  $|S| \geq \rho$  and  $m_i(S) = 0$  otherwise.

Since it will be clear from the context which of the two types of societies we are studying, we do not reflect the type of society in the notation for the material payoffs. We incorporate positional concerns through the following notion of *relative payoff*.

**Definition 46** (Relative Payoff). Let  $(N, \lambda, \rho)$  be a society. The *relative payoff* of individual  $i \in N$  in coalition  $S \in \mathcal{N}$  equals  $r_i(S) = m_i(S) - \bar{u}(S)$ .

The relative payoff measures the material payoff of individual  $i$  relative to the average payoff in the individual's coalition. Note that both for competitive and egalitarian societies, the partition structure uniquely determines the payoffs and thereby the relative payoffs. Note also that in an egalitarian society, the relative payoff is equal to zero for each individual.

For every  $x \in X$  and  $i \in N$ , we denote by  $m_i(x)$  the material payoff of individual  $i$  in state  $x$ , i.e.,  $m_i(x) = m_i(S_i(x))$ , where  $S_i(x)$  is the coalition to which individual  $i$  belongs in state  $x$ . Similarly, we write  $r_i(x) = r_i(S_i(x))$  for the relative payoff of individual  $i$  in state  $x$ .

We are now ready to define the preferences of individual  $i \in N$  over the state space. For every  $x, y \in X$ , we define *relative payoff preferences*  $\geq_i$  over  $X$  by setting  $x \geq_i y$  if and only if either  $m_i(x) > m_i(y)$  or both  $m_i(x) = m_i(y)$  and  $r_i(x) \geq r_i(y)$ . Thus, individual  $i$  prefers the state with the higher material payoff and if two states deliver the same material payoff, the one with the higher relative payoff. We denote the asymmetric part of  $\geq_i$  by  $>_i$ .

The final building block of our social environment is an *effectivity correspondence*  $E$ . For each pair of states  $x, y \in X$ , the effectivity correspondence  $E$  associates a collection of coalitions  $E(x, y) \subseteq \mathcal{N}$  which can move from  $x$  to  $y$ . If  $E(x, y) = \emptyset$ , then no coalition can move from  $x$  to  $y$ . If  $T \in E(x, y)$ , we say that coalition  $T$  can move from state  $x$  to state  $y$ .

We distinguish three groups that are affected by a move of  $T$  from  $x$  to  $y$ : the moving individuals  $T$  themselves, the residual individuals, and the unaffected individuals. The moving individuals  $T$  create a new coalition and leave their former coalition members, the residual individuals, behind. Formally, for every  $x, y \in X$  and  $T \in E(x, y)$ , the residual individuals are given by

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

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

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

The members of the coalitions in  $\mathcal{U}(x, T)$  are called the unaffected individuals and are collected in the set  $U(x, T)$ .

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

individuals stay together.<sup>4</sup> We also assume that the deviating coalition does not have the power to influence the unaffected individuals. The effectivity correspondence is therefore defined as

$$E(x, y) = \{T \in \mathcal{N} \mid T \in y, \mathcal{U}(x, T) \subseteq y, \forall i \in R(x, T), S_i(x) \setminus T \in y\}, \quad x, y \in X.$$

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

It follows from the definition that, for every  $x \in X$ ,  $E(x, x) = x$ , that is, any coalition in  $x$  can choose to stay at  $x$ . Proposition 47 shows that for every  $x, y \in X$  such that  $x \neq y$ , at most two coalitions can be part of  $E(x, y)$ .

**Proposition 47.** *Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. For all  $x, y \in X$  such that  $x \neq y$ , we have  $|E(x, y)| \in \{0, 1, 2\}$ .*

*Proof.* We distinguish two cases: either the moving individuals were part of one coalition or they were part of two or more coalitions before moving.

**Case 1:** For all  $T \in y$ , there exists  $S \in x$  such that  $T \subseteq S$ .

If  $E(x, y) = \emptyset$ , we are done, so assume  $E(x, y) \neq \emptyset$ . Since  $x \neq y$ , there exists  $T \in y$  and  $S \in x$  such that  $T \subsetneq S$ . Because we are in Case 1 and  $E(x, y) \neq \emptyset$ , the coalition  $S$  is uniquely determined. Observe that  $T$  cannot be an unaffected coalition. Hence,  $T$  can only be a moving coalition or a coalition consisting of residual individuals. In the latter case,  $S \setminus T$  is equal to the moving coalition. It follows that  $E(x, y) = \{T, S \setminus T\}$ .

**Case 2:** There exists  $T \in y$  such that, for all  $S \in x$ , we have  $T \setminus S \neq \emptyset$ .

If  $E(x, y) = \emptyset$ , we are done, so assume  $E(x, y) \neq \emptyset$ . Since  $T$  cannot be an unaffected or a residual coalition, the only way to go from  $x$  to  $y$  is by a move of  $T$ . It follows that  $E(x, y) = \{T\}$ .  $\square$

### 4.3 Stability

In what follows, we define two related notions, the Core and the Myopic Stable Set. Both rely on a particular dominance notion. To define it, we write  $y(x, S) \in X$  to denote the state that results from  $E$  when the current state is  $x$  and the moving coalition is  $S$ .

**Definition 48 (Strict Dominance).** Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. A state  $y \in X$  strictly dominates  $x \in X$  by  $S \in \mathcal{N}$  if  $y = y(x, S)$  and  $y \succ_i x$  for all  $i \in S$ .

Let a coalition  $S \in \mathcal{N}$  be given. Let  $f_S : X \rightarrow X$  be the coalitional dominance correspondence, where, for every  $x \in X$ ,  $f_S(x)$  denotes the set of states that strictly dominate  $x$  by  $S$ . Note that  $f_S(x)$  is either the singleton  $\{y(x, S)\}$  or the empty set. We write  $f(x) = \bigcup_{S \in \mathcal{N}} f_S(x)$  to denote the subset of  $X$  consisting of all states that strictly dominate  $x$ . Such states are also referred to as myopic improvements upon  $x$ . The Core equals the set of states which are not strictly dominated by another state.

<sup>4</sup>An alternative specification is the  $\gamma$ -model, which prescribes that residual individuals fall apart into singleton coalitions. This would lead to similar results.

**Definition 49 (Core).** Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. The *Core* equals the set of states  $C = \{x \in X \mid f(x) = \emptyset\}$ .

To define the Myopic Stable Set, it is convenient to define  $\tilde{f}(x) = f(x) \cup \{x\}$ , so  $\tilde{f}(x)$  is obtained from  $f(x)$  by adding the state  $x$ . We define the two-fold composition of  $\tilde{f}$  by

$$\tilde{f}^2(x) = \{z \in X \mid \exists y \in \tilde{f}(x) : z \in \tilde{f}(y)\}.$$

For  $k \in \mathbb{N}$ , we define the  $k$ -fold iteration  $\tilde{f}^k(x)$  by induction as

$$\tilde{f}^k(x) = \{z \in X \mid \exists y \in \tilde{f}^{k-1}(x) : z \in \tilde{f}(y)\}.$$

We define the set of all states that can be reached from  $x$  by a finite number of dominations by

$$\tilde{f}^{\mathbb{N}}(x) = \bigcup_{k \in \mathbb{N}} \tilde{f}^k(x).$$

Because the set  $X$  is finite, there is  $k' \in \mathbb{N}$  such that, for every  $k \geq k'$ ,  $\tilde{f}^{\mathbb{N}}(x) = \tilde{f}^k(x)$ .

We are now ready to introduce the Myopic Stable Set, following the definition of Demuynck, Herings, Saulle, and Seel (2018) for general social environments.<sup>5</sup>

**Definition 50 (Myopic Stable Set).** Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a society. The set  $M \subseteq X$  is a *Myopic Stable Set* if it satisfies the following three conditions:

1. **Deterrence of external deviations:** For every state  $x \in M$ ,  $f(x) \subseteq M$ .
2. **Iterated external stability:** For every  $y \in X \setminus M$ , we have that  $\tilde{f}^{\mathbb{N}}(y) \cap M \neq \emptyset$ .
3. **Minimality:** There is no set  $M' \subsetneq M$  that satisfies Conditions 1 and 2.

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

A Myopic Stable Set is based on a myopic notion of dominance as agents or coalitions do not predict how their decision to change the current state will lead to further changes by other agents or coalitions. Such a notion is natural in complex social environments where the number of possible states is large and agents have little information about the possible actions other agents may take or the incentives of other agents. As shown in Demuynck, Herings, Saulle, and Seel (2018), for finite state spaces there is an equivalence between the set of recurrent states of any dynamic process that selects all myopic improvements with positive probability and the Myopic Stable Set.

In the remainder of this section, we provide some auxiliary results which will be helpful in establishing the main results in the next section. The first set of results follow directly

<sup>5</sup>Demuynck, Herings, Saulle, and Seel (2018) allow for social environments with an infinite state space. For the finite case, their closedness requirement is automatically satisfied and  $\tilde{f}^{\infty}(y)$ , the closure of  $\tilde{f}^{\mathbb{N}}(y)$ , can be replaced by  $\tilde{f}^{\mathbb{N}}(y)$  in the definition of iterated external stability. Thus, for the finite case, the definition of Myopic Stable Set in Demuynck, Herings, Saulle, and Seel (2018) is equivalent to Definition 50.

from previous work. More specifically, by Theorem 3.4 in Demuynck, Herings, Saulle, and Seel (2018), if the dominance correspondence  $\tilde{f}$  is lower hemi-continuous then the Myopic Stable Set is unique. Notice that lower hemi-continuity is trivially satisfied for finite state spaces. Thus, the Myopic Stable Set of the social environment induced by a society is always unique. The property of internal external stability guarantees that the Myopic Stable Set is non-empty. Moreover, by Corollary 3.11 in Demuynck, Herings, Saulle, and Seel (2018), the Core is a subset of the Myopic Stable Set. Contrary to the concept of the Core, which only requires that there is no deviation away from one of its elements, the property of iterated external stability guarantees that the Myopic Stable Set can be reached from any initial state.

We establish a simple but useful result for the case in which some individuals have the same productivity level. Intuitively, if a state in which these individuals are in different coalitions is in the Core, then all states in which these individuals swap their coalitions are also in the Core. To define this formally, for given  $x, y \in X$ , let  $b : N \rightarrow N$  be a *productivity preserving permutation*, i.e.,  $b$  is a one-to-one mapping such that  $b(i) = j$  implies  $\lambda_i = \lambda_j$ . Moreover, we define  $b(S) = \{j \in N \mid \exists i \in S \text{ such that } j = b(i)\}$ .

**Definition 51** (Equivalence). Let  $(N, \lambda, \rho)$  be a society. The states  $x$  and  $y$  in  $X$  are *equivalent* if there is a productivity preserving permutation  $b$  such that for all coalitions  $S \in x$  it holds that  $b(S) \in y$ .

Since  $b$  can be the identity, any state  $x$  is equivalent to itself. We denote the equivalence class of a state  $x$  by  $[x]$ . The next result shows if a state is in the core, then all equivalent states are also in the core.

**Proposition 52.** Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be a social environment induced by a competitive or an egalitarian society. If  $x \in C$ , then  $[x] \subseteq C$ .

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

We do not establish the corresponding analogue of Proposition 52 for the Myopic Stable Set at this point, since it does not admit a simple proof. In the next section, however, we show that the predictions of the Myopic Stable Set and the Core coincide. Thus, if the Myopic Stable Set contains a state  $x$ , it also contains all states in  $[x]$ .

## 4.4 Segregation

In this section, we introduce different notions of segregation. In our framework, segregation means that coalitions are formed between individuals with adjacent productivity levels. We first define segregation at the coalitional level.

**Definition 53** (Segregated Coalition). Let  $(N, \lambda, \rho)$  be a society. A coalition  $S \in \mathcal{N}$  is *segregated* if for all  $i, j \in S$  with  $\lambda_i < \lambda_j$  there is no  $k \in N \setminus S$  such that  $\lambda_i < \lambda_k < \lambda_j$ .

To facilitate a formal definition of segregation at the societal level, we denote the set of all unproductive individuals at state  $x$  by  $Z(x) = \{i \in N \mid |S_i(x)| < \rho\}$  and the collection of all productive coalitions at state  $x$  by  $\mathcal{P}(x) = \{S \in x \mid |S| \geq \rho\}$ . We introduce two notions which capture the idea of segregation at the societal level. The first one is called *bottom-up segregation*. A state is bottom-up segregated if it satisfies three properties: *coalitional assortativity*, *homophily*, and *exclusion of the strong*.

**Definition 54** (Bottom-up Segregated State). Let  $(N, \lambda, \rho)$  be a society. The state  $x \in X$  is bottom-up segregated if the following three properties hold:

- i. **Coalitional assortativity:** For every  $S, T \in \mathcal{P}(x)$  with  $S \neq T$ , we have  $\lambda_{w(S)} \geq \lambda_{s(T)}$  or  $\lambda_{w(T)} \geq \lambda_{s(S)}$ .
- ii. **Homophily:** If  $S \in x$  satisfies  $|S| > \rho$ , then, for all  $i, j \in S$ , we have  $\lambda_i = \lambda_j$ .
- iii. **Exclusion of the strong:** It holds that  $|Z(x)| < \rho$ . If  $Z(x) \neq \emptyset$ , then for every  $i \in Z(x)$  and for every  $j \in N \setminus Z(x)$ , we have  $\lambda_i \geq \lambda_j$ .

Coalitional assortativity requires that given any two productive coalitions, a weak individual in one coalition has weakly higher productivity than a strong individual in the other coalition. Homophily states that if a coalition strictly exceeds the productivity threshold, then all individuals in the coalition have the same productivity. Exclusion of the strong means that only the strongest individuals may not be part of a productive coalition.

We denote the set of all bottom-up segregated states by  $\underline{X}$ .

**Example 9.** Consider a society  $(N, \lambda, \rho)$  with  $N = \{1, \dots, 7\}$ , productivities given by  $\lambda_1 = \dots = \lambda_3 > \lambda_4 = \dots = \lambda_7$ , and a productivity threshold  $\rho = 2$ . The state  $x = \{\{1, 2\}, \{3\}, \{4, 5, 6, 7\}\}$  is an example of a bottom-up segregated state. Since all individuals in a coalition have the same productivity, coalitional assortativity and homophily are trivially satisfied. Exclusion of the strong holds since the only member of an unproductive coalition is the strong individual 3.

The following lemma shows that productive coalitions in a bottom-up segregated state are indeed segregated.

**Lemma 55.** Let  $(N, \lambda, \rho)$  be a society. If  $x \in \underline{X}$  and  $S \in \mathcal{P}(x)$ , then  $S$  is segregated.

*Proof.* Suppose not. Let  $x \in \underline{X}$ ,  $S \in \mathcal{P}(x)$ ,  $i, j \in S$ , and  $k \in N \setminus S$  be such that  $\lambda_i < \lambda_k < \lambda_j$ . Let  $T \in x$  be such that  $k \in T$ . By homophily, it holds that  $|S| = \rho$ . If  $T \in \mathcal{P}(x)$ , then, by coalitional assortativity,  $\lambda_{w(S)} \geq \lambda_{s(T)} \geq \lambda_k$  or  $\lambda_k \geq \lambda_{w(T)} \geq \lambda_{s(S)}$ , which contradicts  $i, j \in S$  and  $\lambda_i < \lambda_k < \lambda_j$ . If  $T \notin \mathcal{P}(x)$ , then  $k \in Z(x)$ , so by exclusion of the strong we have  $\lambda_k \geq \lambda_{s(S)}$ , leading to a contradiction with  $j \in S$  and  $\lambda_k < \lambda_j$ .  $\square$

In what follows, we provide an algorithmic procedure to find all bottom-up segregated states of a given society. The procedure results in a weighted directed rooted tree  $(V, A, v^0)$ , where  $V$  is a set of decision nodes,  $A \subseteq V \times V$  is a set of ordered pairs of nodes called arcs, and  $v^0$  is a given initial decision node. To each arc  $a \in A$ , we associate an integer  $\mu_a \in \mathbb{N}$  called its weight. This weight  $\mu_a$  corresponds to the size of the coalition that is formed at



the current decision node. Given some  $v \in V$ , we denote by  $A^v \subseteq A$  the set of all arcs on the directed path from  $v^0$  to  $v$ . Since  $(V, A, v^0)$  is a directed rooted tree, the set  $A^v$  is uniquely determined. To each decision node  $v$ , we associate a number  $n(v)$  corresponding to the number of individuals that are not yet assigned to a coalition.

**Procedure 56.** Let  $(N, \lambda, \rho)$  be a society. Let  $v = v^0$  with  $v^0$  a given decision node. The initial set of arcs  $A$  is equal to the empty set.

**Step 1** Define  $n(v) = n - \sum_{a \in A^v} \mu_a$ . If  $n(v) \geq \rho$ , then add

$$k = \max \{1, n(v) - \min \{i \in \{1, \dots, n(v)\} \mid \lambda_i = \lambda_{n(v)}\} - \rho + 2\}$$

arcs starting at  $v$  and ending at distinct new decision nodes. The arcs all have different weights, ranging from  $\rho$  to  $\rho + k - 1$ . If  $0 \leq n(v) < \rho$ , then  $v$  is a terminal decision node and no new decision node is created.

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

**Step 3** Consider the weighted rooted directed tree  $(V, A, v^0)$  resulting from Steps 1 and 2. For every terminal node  $v$ , let  $(a_1, \dots, a_\ell)$  be the arcs on the path from  $v^0$  to  $v$ , and, for every partition  $P \in \mathbb{P}\{n - \sum_{k=1}^{\ell} \mu_{a_k}, \dots, 1\}$ , generate the state

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

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

Intuitively, Procedure 56 works as follows. First, the highest numbered individuals form a productive coalition that satisfies homogeneity. Then the remaining highest numbered individuals form the next productive coalition that satisfies homogeneity, and so on. If some individuals are not part of a productive coalition of this bottom-up process, then those are the lowest numbered ones, i.e., the most productive ones, and they split up in an arbitrary way. The final step of Procedure 56 picks up equivalent states.

We use Example 9 to illustrate Procedure 56. The weighted directed rooted tree generated in Steps 1 and 2 is illustrated in Figure 1. This directed rooted tree has five terminal nodes, resulting in five states in Step 3:

$$\begin{aligned} x_1 &= \{\{1, 2, 3\}, \{4, 5\}, \{6, 7\}\}, \\ x_2 &= \{\{1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}\}, \\ x_3 &= \{\{1, 2\}, \{3, 4\}, \{5, 6, 7\}\}, \\ x_4 &= \{\{1\}, \{2, 3\}, \{4, 5, 6, 7\}\}, \\ x_5 &= \{\{1, 2, 3\}, \{4, 5, 6, 7\}\}. \end{aligned}$$

Step 4 adds equivalent states. By elementary combinatorics, this leads to two additional states for  $x_1$ , eight additional states for  $x_2$ , eleven additional states for  $x_3$ , two additional states for  $x_4$ , and no additional states for  $x_5$ . None of these states is identical. Thus, we have a total of 28 bottom-up segregated states. We refer to the appendix for a full enumeration of the bottom-up segregated states.

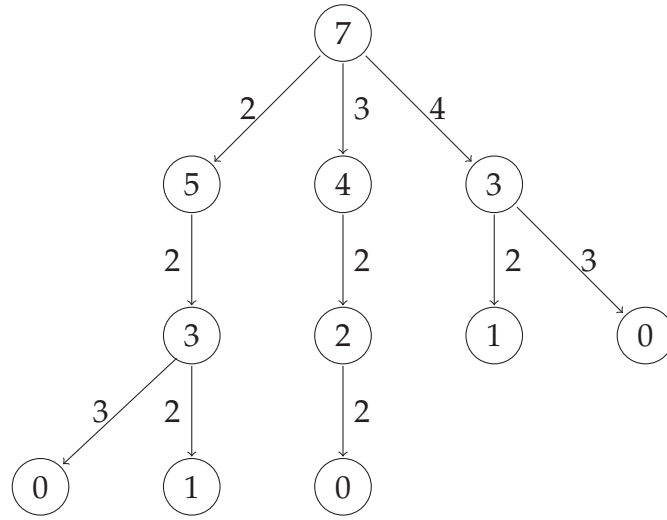


Figure 4.1: Result of Steps 1 and 2 of Procedure 56 for Example 9.

**Proposition 57.** Let  $(N, \lambda, \rho)$  be a society. Procedure 56 generates the set of bottom-up segregated states  $\underline{X}$ .

*Proof.* See Appendix. □

The second notion of societal segregation is *top-down segregation*. As bottom-up segregation, it requires three properties. The first two properties, coalitional assortativity and homophily, are identical as for bottom-down segregation. The third one replaces exclusion of the strong by exclusion of the weak, i.e., the least productive individuals can be excluded.

**Definition 58 (Top-down Segregated State).** Let  $(N, \lambda, \rho)$  be a society. The state  $x \in X$  is top-down segregated if the following three properties hold:

- i. **Coalitional assortativity:** For every  $S, T \in \mathcal{P}(x)$  with  $S \neq T$ , we have  $\lambda_{w(S)} \geq \lambda_{s(T)}$  or  $\lambda_{w(T)} \geq \lambda_{s(S)}$ .
- ii. **Homophily:** If  $S \in x$  satisfies  $|S| > \rho$ , then, for all  $i, j \in S$ , we have  $\lambda_i = \lambda_j$ .
- iii. **Exclusion of the weak:** It holds that  $|Z(x)| < \rho$ . If  $Z(x) \neq \emptyset$ , then for every  $i \in Z(x)$  and for every  $j \in N \setminus Z(x)$ , we have  $\lambda_i \leq \lambda_j$ .

We denote the set of all top-down segregated states by  $\overline{X}$ . The following result states that every productive coalition in a top-down segregated state is segregated. The proof is analogous to Lemma 55 and therefore omitted.

**Lemma 59.** Let  $(N, \lambda, \rho)$  be a society. If  $x \in \overline{X}$  and  $S \in \mathcal{P}(x)$ , then  $S$  is segregated.

The procedure to find all top-down segregated states also proceeds along the same lines as for bottom-down segregated states.

**Procedure 60.** Let  $(N, \lambda, \rho)$  be a society. Let  $v = v^0$  with  $v^0$  a given decision node. The initial set of arcs  $A$  is equal to the empty set.

**Step 1** Define  $n(v) = n - \sum_{a \in Av} \mu_a$ . If  $n(v) \geq \rho$ , then add

$$k = \max \{1, n(v) - \min \{i \in \{1, \dots, n(v)\} \mid \lambda_i = \lambda_{n(v)}\} - \rho + 2\}$$

arcs starting at  $v$  and ending at distinct new decision nodes. The arcs all have different weights, ranging from  $\rho$  to  $\rho + k - 1$ . If  $0 \leq n(v) < \rho$ , then  $v$  is a terminal decision node and no new decision node is created.

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

**Step 3** Consider the weighted rooted directed tree  $(V, A, v^0)$  resulting from Steps 1 and 2. For every terminal node  $v$ , let  $(a_1, \dots, a_\ell)$  be the arcs on the path from  $v^0$  to  $v$ , and, for every partition  $P \in \mathbb{P}\{\sum_{k=1}^{\ell} \mu_{a_k} + 1, \dots, n\}$  generate the state

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

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

**Proposition 61.** *Let  $(N, \lambda, \rho)$  be a society. Procedure 60 generates the set of top-down segregated states  $\bar{X}$ .*

As the procedure is the mirror image of the case for bottom-up segregation, we omit the proof that Procedure 60 results in the set of top-down segregated states.

## 4.5 Results for Competitive and Egalitarian Societies

In this section, we study the general properties of competitive and egalitarian societies with relative payoff preferences.

The following result is the main characterization result for competitive societies.

**Theorem 62.** *Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be the social environment induced by a competitive society  $(N, \lambda, \rho)$ . A state is in the Core of  $\Gamma$  if and only if it is a bottom-up segregated state. Moreover, the Core coincides with the Myopic Stable Set.*

*Proof.* See Appendix. □

Both the Core and the Myopic Stable Set predict the set of bottom-up segregated states. Let us start with an intuition for the Core. An existing productive coalition has two different ways of improving its payoffs. Either it can get rid off a member with above average productivity or it can add less productive individuals to the coalition. When a state satisfies homophily, the first option does not work as any resulting coalition would be unproductive. The second option is prevented when a state satisfies coalitional assortativity as the less productive individuals would not want to switch to a more productive coalition. Finally, if the total number of individuals in unproductive coalitions is below  $\rho$ , an unproductive individual can only form coalitions with already productive individuals. However, if the unproductive individuals have a high productivity as a consequence of

exclusion of the strong, other individuals will not form coalitions with them due to the resulting lower social rank.

The equivalence of the Core and the Myopic Stable Set gives an additional dynamic foundation for the prediction: from any state that is not bottom-up segregated, there is a path of myopic improvements to a bottom-up segregated state. In fact, any dynamic process that selects myopic improvements with positive probability will have the Myopic Stable Set as the set of recurrent states.

We now turn to the special case where all individuals differ in productivity.

**Definition 63** (Complete Heterogeneity). A society  $(N, \lambda, \rho)$  is *completely heterogeneous* if all individuals differ in their productivity, i.e., for all  $i, j \in N$  with  $i \neq j$ , we have  $\lambda_i \neq \lambda_j$ .

Let  $\ell$  and  $o$  be non-negative integers such that  $o < \rho$  and the number of individuals in the society equals  $n = \ell\rho + o$ . The number  $\ell$  is equal to the highest possible number of productive coalitions. For  $k \in \{1, \dots, \ell\}$ , we define  $B^k = \{n - \rho k + 1, \dots, n - (k - 1)\rho\}$  and call it the  $k^{\text{th}}$  bottom-up segregated coalition. These coalitions are collected in the set  $\mathcal{B} = \{B^k \mid k = 1, \dots, \ell\}$ .

**Proposition 64.** Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be the social environment induced by the completely heterogeneous competitive society  $(N, \lambda, \rho)$ . Then the Core consists of those states which contain the collection  $\mathcal{B}$ , i.e.,  $C = \{x \in X \mid \mathcal{B} \subseteq x\}$ .

*Proof.* By Theorem 62, the Core coincides with the set of bottom-up segregated states  $\underline{X}$ . By Proposition 57, Procedure 56 generates the entire set  $\underline{X}$ . Using the fact that the society is completely heterogeneous, it follows that in Step 1 of the procedure only arcs with weight  $\rho$  are generated, leading to exactly the states in  $\{x \in X \mid \mathcal{B} \subseteq x\}$  in Step 3. By the fact that the society is completely heterogeneous, there are no equivalent states  $x$  and  $y$  such that  $x \neq y$ . Therefore, no new states are added in Step 4 of Procedure 56.  $\square$

The previous proposition uniquely characterizes the productive coalitions in the Core. It allows for an arbitrary partition of the unproductive individuals. Their material payoff is zero in either case and thus their relative payoff is also zero.

If  $n$  is an integer multiple of  $\rho$ , then the Core is unique and it coincides with the collection of bottom-up segregated coalitions  $\mathcal{B}$ . Similarly, if  $n = \rho\ell + 1$ , then the Core is unique and it coincides with the collection of bottom-up segregated coalitions  $\mathcal{B}$  and the singleton  $\{1\}$ .

We now turn to the analysis of egalitarian societies. Recall that relative payoffs are now zero for all individuals. Thus, the preferences coincide with the preferences over the material payoffs. The following proposition is the main characterization result for egalitarian societies.

**Theorem 65.** Let  $\Gamma = (N, X, E, \{\succeq_i\}_{i \in N})$  be the social environment induced by an egalitarian society  $(N, \lambda, \rho)$ . A state is in the Core of  $\Gamma$  if and only if it is a top-down segregated state. Moreover, the Core coincides with the Myopic Stable Set.

*Proof.* See Appendix.  $\square$

Contrary to the case of competitive societies, in egalitarian societies members of productive coalitions prefer to have more productive individuals in their coalition.

Let  $\ell$  and  $o$  be non-negative integers such that  $o < \rho$  and the number of individuals in the society equals  $n = \ell\rho + o$ . For  $k \in \{1, \dots, \ell\}$ , we define  $T^k = \{\rho k - \rho + 1, \dots, \rho k\}$  and call it the  $k^{\text{th}}$  top-down segregated coalition. These coalitions are collected in the set  $\mathcal{T} = \{T^k \mid k = 1, \dots, \ell\}$ .

**Proposition 66.** *Let  $\Gamma = (N, X, E, \{\geq_i\}_{i \in N})$  be the social environment induced by the completely heterogeneous egalitarian society  $(N, \lambda, \rho)$ . Then the Core consists of those states which contain the collection  $\mathcal{T}$ , i.e.,  $C = \{x \in X \mid \mathcal{T} \subseteq x\}$ .*

*Proof.* By Theorem 65, the Core coincides with the set of top-down segregated states  $\bar{X}$ . Procedure 60 generates the entire set  $\bar{X}$ . Using the fact that the society is completely heterogeneous, it follows that in Step 1 of the procedure only arcs with weight  $\rho$  are generated, leading to exactly the states in  $\{x \in X \mid \mathcal{T} \subseteq x\}$  in Step 3. By the fact that the society is completely heterogeneous, there are no equivalent states  $x$  and  $y$  such that  $x \neq y$ . Therefore, no new states are added in Step 4 of Procedure 60.  $\square$

## 4.6 Material Efficiency

In this section, we compare competitive and egalitarian allocations in terms of the resulting material efficiency.

**Definition 67** (Material Efficiency). Let  $(N, \lambda, \rho)$  be a society. The *material efficiency* of a state  $x \in X$  is equal to  $\sum_{i \in N} m_i(x)$ .

The maximal material efficiency a state  $x \in X$  can reach is given by  $\sum_{i \in N} \lambda_i$ . We now compare competitive and egalitarian societies in terms of the resulting material efficiency.

**Theorem 68.** *Let  $(N, \lambda, \rho)$  be a completely heterogeneous society. Every Core element of the social environment induced by the egalitarian society is at least as materially efficient as every Core element of the social environment induced by the competitive society. They are equally materially efficient if and only if  $n$  is an integer multiple of  $\rho$ .*

*Proof.* Recall that in the competitive case the Core is equal to  $\underline{X} = \{x \in X \mid \mathcal{B} \subseteq x\}$  by Proposition 64 and in the egalitarian case it equals  $\bar{X} = \{x \in X \mid \mathcal{T} \subseteq x\}$  by Proposition 66.

Let some  $\underline{x} \in \underline{X}$  and some  $\bar{x} \in \bar{X}$  be given. We have that  $\sum_{i \in N} m_i(\underline{x}) = \sum_{i \in N} \lambda_i - \sum_{i \in Z(\underline{x})} \lambda_i$  and  $\sum_{i \in N} m_i(\bar{x}) = \sum_{i \in N} \lambda_i - \sum_{i \in Z(\bar{x})} \lambda_i$ .

Let  $\ell$  and  $o$  be non-negative integers such that  $o < \rho$  and the number of individuals in the society equals  $n = \ell\rho + o$ . As the society is completely heterogeneous, it holds that  $\sum_{i \in Z(\underline{x})} \lambda_i = (\lambda_1 + \dots + \lambda_o) \geq (\lambda_{n-o+1} + \dots + \lambda_n) = \sum_{i \in Z(\bar{x})} \lambda_i$ , so  $\sum_{i \in N} m_i(\underline{x}) \leq \sum_{i \in N} m_i(\bar{x})$ , where equality holds if and only if  $o = 0$ .  $\square$

## 4.7 Concluding Remarks

We have studied coalition formation among individuals who differ in productivity. The joint output of a coalition is determined by the sum of productivities if a coalition exceeds

a minimal threshold of members. We consider two types of societies, competitive societies in which the surplus of a coalition is split according to productivity and egalitarian societies in which coalitions split their surplus equally. The preferences of coalition members depend on the share of the surplus, but also on the relative payoff in the group.

Our analysis of the societies uses two different stability concepts, the classic static concept of the Core and the Myopic Stable Set, which enables a dynamic interpretation. The two concepts are shown to yield the same predictions. Both competitive and egalitarian societies lead to segregated societies, i.e., coalitions are only formed between individuals with similar productivities.

In contrast to a superficial intuition, the only individuals who might not be part of a productive coalition are the most productive individuals in a competitive society and the least productive individuals in an egalitarian society. To understand the intuition for these predictions, note that members of a competitive society will lose in social rank when forming coalitions with very productive individuals. Thus, the last (least productive individuals) are first in forming coalitions and the first (most productive individuals) last. The most productive individuals might even fail to attract other coalition members at all. On the contrary, in the egalitarian society the most productive individuals are very attractive as coalition members, since other coalition members gain from their productivity.

## 4.8 Appendix

### Bottom-up Segregated States in the Example 9.

$[x_1]$	$\{\{1, 2, 3\}, \{4, 5\}, \{6, 7\}\}$ $\{\{1, 2, 3\}, \{4, 6\}, \{5, 7\}\}$ $\{\{1, 2, 3\}, \{4, 7\}, \{5, 6\}\}$
$[x_2]$	$\{\{1\}, \{2, 3\}, \{4, 5\}, \{6, 7\}\}$ $\{\{1\}, \{2, 3\}, \{4, 6\}, \{5, 7\}\}$ $\{\{1\}, \{2, 3\}, \{4, 7\}, \{5, 6\}\}$ $\{\{2\}, \{1, 3\}, \{4, 5\}, \{6, 7\}\}$ $\{\{2\}, \{1, 3\}, \{4, 6\}, \{5, 7\}\}$ $\{\{2\}, \{1, 3\}, \{4, 7\}, \{5, 6\}\}$ $\{\{3\}, \{1, 2\}, \{4, 5\}, \{6, 7\}\}$ $\{\{3\}, \{1, 2\}, \{4, 6\}, \{5, 7\}\}$ $\{\{3\}, \{1, 2\}, \{4, 7\}, \{5, 6\}\}$
$[x_3]$	$\{\{1, 2\}, \{3, 4\}, \{5, 6, 7\}\}$ $\{\{1, 2\}, \{3, 5\}, \{4, 6, 7\}\}$ $\{\{1, 2\}, \{3, 6\}, \{4, 5, 7\}\}$ $\{\{1, 2\}, \{3, 7\}, \{4, 5, 6\}\}$ $\{\{1, 3\}, \{2, 4\}, \{5, 6, 7\}\}$ $\{\{1, 3\}, \{2, 5\}, \{4, 6, 7\}\}$ $\{\{1, 3\}, \{2, 6\}, \{4, 5, 7\}\}$ $\{\{1, 3\}, \{2, 7\}, \{4, 5, 6\}\}$ $\{\{2, 3\}, \{1, 4\}, \{5, 6, 7\}\}$ $\{\{2, 3\}, \{1, 5\}, \{4, 6, 7\}\}$ $\{\{2, 3\}, \{1, 6\}, \{4, 5, 7\}\}$ $\{\{2, 3\}, \{1, 7\}, \{4, 5, 6\}\}$
$[x_4]$	$\{\{1\}, \{2, 3\}, \{4, 5, 6, 7\}\}$ $\{\{2\}, \{1, 3\}, \{4, 5, 6, 7\}\}$ $\{\{3\}, \{1, 2\}, \{4, 5, 6, 7\}\}$
$[x_5]$	$\{\{1, 2, 3\}, \{4, 5, 6, 7\}\}$

### Proof of Theorem 57

*Proof.* We first show that any state generated by Procedure 56 is bottom-up segregated.

We start with a state generated in Step 3 from an arc  $(a_1, \dots, a_\ell)$ . Since all productive coalitions consist of individuals with adjacent subscripts and hence adjacent productivities, we have coalitional assortativity. If, for some  $k \in \{1, \dots, \ell\}$ ,  $\mu_{a_k} > \rho$ , then it follows from the selection of the number of added arcs in Step 1 that all individuals involved in the associated coalition have the same productivity, so homophily is satisfied. Finally, the individuals in the set  $\{n - \sum_{k=1}^{\ell} \mu_{a_k}, \dots, 1\}$  cannot form a productive coalition themselves and have weakly higher productivity than individuals in the productive coalitions, so exclusion of the strong is satisfied. To complete the first part of the proof, note that all three properties of a bottom-up segregated state are preserved under productivity preserving

permutations, so the equivalent states generated in Step 4 are bottom-up segregated as well.

It remains to be shown that all bottom-up segregated states are generated by Procedure 56. To do so, we show that for each bottom-up segregated state, there is one permutation of a state generated in Steps 1-3. The key is to define the permutation/relabeling in a precise way.

Let  $x = \{S^1, \dots, S^m\} \in \underline{X}$  be a bottom-up segregated state with productive coalitions  $\mathcal{P}(x) = \{S^1, \dots, S^\ell\}$ , where  $S^1, \dots, S^\ell$  are chosen such that

$$\lambda_{s(S^1)} \leq \lambda_{w(S^2)} \leq \dots \leq \lambda_{s(S^{\ell-1})} \leq \lambda_{w(S^\ell)},$$

and unproductive coalitions  $S^{\ell+1}, \dots, S^m$ . For  $k \in \{1, \dots, m\}$ , we denote the cardinality  $|S^k|$  of coalition  $S^k$  by  $p_k$ .

We define the permutation  $b : N \rightarrow N$  as follows. Let some  $i \in N$  be given. The number  $k' \in \{1, \dots, \ell + 1\}$  is uniquely determined in the following way. If

$$i \in \{n - \sum_{k=1}^{\ell} p_k, \dots, 1\},$$

then  $k' = \ell + 1$ . Otherwise,  $k' \in \{1, \dots, \ell\}$  is chosen such that

$$i \in \{n - \sum_{k=1}^{k'-1} p_k, \dots, n - \sum_{k=1}^{k'} p_k + 1\}.$$

In case  $k' \leq \ell$ , we write  $S^{k'} = \{j_1, \dots, j_{p_{k'}}\}$  with  $j_1 > \dots > j_{p_{k'}}$  and define

$$b(i) = j_{n - \sum_{k=1}^{k'-1} p_k - i + 1}.$$

In case  $k' = \ell + 1$ , we write  $S^{\ell+1} \cup \dots \cup S^m = \{j_1, \dots, j_{n - \sum_{k=1}^{\ell} p_k}\}$  with  $j_1 > \dots > j_{n - \sum_{k=1}^{\ell} p_k}$  and define

$$b(i) = j_{n - \sum_{k=1}^{\ell} p_k - i + 1}.$$

One of the directed paths  $(a_1, \dots, a_\ell)$  generated by Procedure 56 is such that, for  $k = 1, \dots, \ell$ ,  $\mu_{a_k} = p_k$ . Let the permutation  $P \in \mathbb{P}\{n - \sum_{k=1}^{\ell} \mu_{a_k}, \dots, 1\}$  be defined by  $P = \{b^{-1}(S^k) \mid k \in \{\ell + 1, \dots, m\}\}$ . In Step 3, the directed path  $(a_1, \dots, a_\ell)$  together with the permutation  $P$  generates the state

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

Clearly, for all  $S \in y$  it holds that  $b(S) \in x$ . It remains to be shown that  $b$  is productivity preserving, so  $b(i) = j$  implies  $\lambda_i = \lambda_j$ , in order to conclude that  $x \in [y]$ , so is generated in Step 4.

Suppose there is  $i, j \in N$  such that  $b(i) = j$  and  $\lambda_i \neq \lambda_j$ . Let  $i'$  be the highest numbered individual with this property and define  $j' = b(i')$ . Since  $x$  satisfies exclusion of the strong and by the construction of  $b$ , it cannot hold that  $j' \in S^{\ell+1} \cup \dots \cup S^m$ . Let  $k' \leq \ell$  be such that  $j' \in S^{k'}$ . Since  $i'$  is the highest numbered individual among  $i \in N$  such that  $\lambda_i \neq \lambda_{b(i)}$ , it must hold that  $\lambda_{i'} < \lambda_{j'}$ . It follows that there is  $j > i'$  such that  $j \notin S^1 \cup \dots \cup S^{k'}$ . Suppose there is  $k \in \{k' + 1, \dots, \ell\}$  such that  $j \in S^k$ . Since  $\lambda_{w(S^k)} \leq \lambda_j < \lambda_{j'} \leq \lambda_{s(S^{k'})}$ , we obtain a contradiction to  $x$  satisfying coalitional assortativity. Consequently, it holds that  $j \in S^{\ell+1} \cup \dots \cup S^m$ . However, in that case  $x$  violates exclusion of the strong since  $\lambda_j \leq \lambda_{i'} < \lambda_{j'}$ . We have obtained a contradiction. Consequently,  $b(i) = j$  implies  $\lambda_i = \lambda_j$ .  $\square$



## Proof of Theorem 62

The following lemma is helpful in the proof of Theorem 62.

**Lemma 69.** *Let  $\Gamma = (N, X, E, \{\geq_i\}_{i \in N})$  be a social environment induced by a competitive society  $(N, \lambda, \rho)$ . Let  $x \in X$  and  $S \in x$  be such that  $|S| > \rho$  and  $\lambda_i \neq \lambda_j$  for some  $i, j \in S$ . Then  $f_{S \setminus \{s(S)\}}(x) \neq \emptyset$ .*

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

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

**Proof of Theorem 62.** The proof consists of three parts. We show first that  $\underline{X} \subseteq C$ , next that  $\underline{X} \supseteq C$ , thereby showing that the Core coincides with the set of bottom-up segregated states, and finally that the Myopic Stable Set coincides with the Core.

**Part 1.**  $\underline{X} \subseteq C$ .

Let  $\underline{x} \in X$  be a bottom-up segregated state, where  $\underline{x} = \{S^1, \dots, S^m\}$ ,  $\mathcal{P}(\underline{x}) = \{S^1, \dots, S^\ell\}$ , and, for every  $k \in \{1, \dots, \ell - 1\}$ ,  $\lambda_{s(S^k)} \leq \lambda_{w(S^{k+1})}$ . Moreover, we write  $S^1 = \{s_1^1, \dots, s_p^1\}$  with  $s_1^1 < \dots < s_p^1$ .

Towards a contradiction, suppose that  $\underline{x} \notin C$ . Thus, there exists a coalition  $T = \{t_1, \dots, t_q\}$  with  $t_1 < \dots < t_q$  and a state  $y$  such that  $y \in f_T(\underline{x})$ . Notice that any deviating coalition has to be productive, i.e.,  $|T| \geq \rho$ . We proceed by induction.

**Step 1:**  $S^1 \cap T = \emptyset$ .

Notice that  $S^1$  is productive. For all  $i \in S^1 \cap T$ , it holds that  $m_i(S^1) = m_i(T)$ . Then a deviation has to rely on an improvement in relative payoff. We consider two cases: a case in which, for all  $i, j \in S^1$ ,  $\lambda_i = \lambda_j$  and a case in which there are some  $i, j \in S^1$  with  $\lambda_i \neq \lambda_j$ .

In the former case, we have that  $\bar{u}(S^1) = \lambda_n \leq \bar{u}(T)$  as, by assumption,  $n$  is a weak individual in the society. Therefore, for all  $i \in S^1 \cap T$ , it holds that

$$r_i(S^1) = m_i(S^1) - \bar{u}(S^1) \geq m_i(T) - \bar{u}(T) = r_i(T),$$

so it follows that  $S^1 \cap T = \emptyset$ .

In the latter case, homophily of  $\underline{x}$  implies that  $|S^1| = \rho$ . We define the collection of  $\rho$  weakest individuals in  $T$  by  $T' = \{t_{q-\rho+1}, \dots, t_q\}$ . It holds that  $\bar{u}(T') \leq \bar{u}(T)$ . For  $i = 0, \dots, \rho - 1$ , it holds that  $\lambda_{t_{q-i}} \geq \lambda_{s_{p-i}^1}$ , so we have that  $\bar{u}(S^1) \leq \bar{u}(T') \leq \bar{u}(T)$ . Therefore, for every  $i \in S^1 \cap T$ , it holds that

$$r_i(S^1) = m_i(S^1) - \bar{u}(S^1) \geq m_i(T) - \bar{u}(T) = r_i(T),$$

so it follows that  $S^1 \cap T = \emptyset$ .

**Step 2:** If  $\bigcup_{\kappa=1}^k S^\kappa \cap T = \emptyset$ , then  $S^{k+1} \cap T = \emptyset$ .

If  $k \geq \ell$ , then  $S^{k+1} \notin \mathcal{P}(\underline{x})$ . As  $\bigcup_{\kappa=1}^k S^\kappa \cap T = \emptyset$  and  $|Z(\underline{x})| < \rho$  it follows that  $|T| < \rho$ , a contradiction to  $|T| \geq \rho$ .

If  $k < \ell$ , then  $S^{k+1} \in \mathcal{P}(\underline{x})$  and we write  $S^{k+1} = \{s_1^{k+1}, \dots, s_p^{k+1}\}$  with  $s_1^{k+1} < \dots < s_p^{k+1}$ . For every  $i \in S^{k+1} \cap T$ , it holds that  $m_i(S^{k+1}) = m_i(T)$  since  $|T| \geq \rho$ . A deviation has to rely on an improvement in relative payoff. If for all  $i, j \in S^{k+1}$ ,  $\lambda_i = \lambda_j$ , then  $\bar{u}(S^{k+1}) =$

$\lambda_{s_p^{k+1}} \leq \bar{u}(T)$  as, by assumption,  $s_p^{k+1}$  is a weak individual in  $N \setminus \bigcup_{\kappa=1}^k S^\kappa$ . Therefore, for all  $i \in S^{k+1} \cap T$ , it holds that

$$r_i(S^{k+1}) = m_i(S^{k+1}) - \bar{u}(S^{k+1}) \geq m_i(T) - \bar{u}(T) = r_i(T),$$

so  $S^{k+1} \cap T = \emptyset$ . If there is some  $i, j \in S^{k+1}$  with  $\lambda_i \neq \lambda_j$ , by homophily,  $|S^{k+1}| = \rho$ . We define  $T' = \{t_{q-\rho+1}, \dots, t_q\}$ , so  $\bar{u}(T') \leq \bar{u}(T)$ . Since  $\lambda_{t_{q-i}} \geq \lambda_{s_{p-i}^{k+1}}$  for  $i = 0, \dots, \rho - 1$ , it follows that  $\bar{u}(S^{k+1}) \leq \bar{u}(T') \leq \bar{u}(T)$ . Therefore, for every  $i \in S^{k+1} \cap T$ , it holds that

$$r_i(S^{k+1}) = m_i(S^{k+1}) - \bar{u}(S^{k+1}) \geq m_i(T) - \bar{u}(T) = r_i(T),$$

so  $S^{k+1} \cap T = \emptyset$ .

By Steps 1 and 2 it follows that  $T = \emptyset$  and we have obtained a contradiction. Consequently, it holds that  $\underline{x} \in C$ .

**Part 2.**  $\underline{X} \supseteq C$ .

Let some  $y \in C$  be given. Towards a contradiction, suppose that  $y$  is not bottom-up segregated.

If homophily is violated, then there exists a coalition  $S \in y$  with  $|S| > \rho$  such that for some  $i, j \in S$ ,  $\lambda_i > \lambda_j$ . Thus, by Lemma 69, it holds that  $f_{S \setminus \{s(S)\}}(y) \neq \emptyset$ , a contradiction to  $y$  being in the Core.

If coalitional assortativity is violated, then there are coalitions  $S, T \in \mathcal{P}(y)$  such that  $\lambda_{w(S)} < \lambda_{s(T)}$  and  $\lambda_{w(T)} < \lambda_{s(S)}$ . We write  $S = \{s_1, \dots, s_p\}$  and  $T = \{t_1, \dots, t_q\}$  with  $s_1 < \dots < s_p$  and  $t_1 < \dots < t_q$ . We show that for  $S' = \{t_q, s_2, \dots, s_p\}$  it holds that  $f_{S'}(y) \neq \emptyset$  or for  $T' = \{s_p, t_2, \dots, t_q\}$  it holds that  $f_{T'}(y) \neq \emptyset$ .

Since  $S$  and  $S'$  are productive, we have  $m_i(S) = m_i(S')$  for all  $i \in S \cap S'$ . Moreover, since  $|S| = |S'|$  and  $\lambda_{s_1} > \lambda_{t_q}$ , we have

$$\bar{u}(S) = \frac{\lambda_{s_1} + \sum_{i=2}^p \lambda_{s_i}}{|S|} > \frac{\lambda_{t_q} + \sum_{i=2}^p \lambda_{s_i}}{|S'|} = \bar{u}(S').$$

Thus, for all  $i \in S \cap S' = S' \setminus \{t_q\}$ , it holds that  $r_i(S) = m_i(S) - \bar{u}(S) < m_i(S') - \bar{u}(S') = r_i(S')$ . By symmetry, it also holds for all  $i \in T \cap T' = T' \setminus \{s_p\}$  that  $m_i(T) = m_i(T')$  and  $r_i(T) < r_i(T')$ .

Towards a contradiction, suppose that  $s_p = w(S)$  has no incentive to participate in  $T'$  and  $t_q = w(T)$  has no incentive to form  $S'$ , so  $r_{w(S)}(S) \geq r_{w(S)}(T')$  and  $r_{w(T)}(T) \geq r_{w(T)}(S')$ . Thus, by definition,

$$\begin{aligned} m_{w(S)}(S) - \frac{\sum_{i \in S} \lambda_i}{p} &\geq m_{w(S)}(T') - \frac{\lambda_{w(S)}}{q} - \frac{\sum_{i \in T \setminus \{s(T)\}} \lambda_i}{q}, \\ m_{w(T)}(T) - \frac{\sum_{i \in T} \lambda_i}{q} &\geq m_{w(T)}(S') - \frac{\lambda_{w(T)}}{p} - \frac{\sum_{i \in S \setminus \{s(S)\}} \lambda_i}{p}. \end{aligned}$$

These expressions can be simplified to get

$$\begin{aligned} p\lambda_{w(S)} + p \sum_{i \in T \setminus \{s(T)\}} \lambda_i &\geq q \sum_{i \in S} \lambda_i, \\ q\lambda_{w(T)} + q \sum_{i \in S \setminus \{s(S)\}} \lambda_i &\geq p \sum_{i \in T} \lambda_i. \end{aligned}$$

Adding up these two inequalities and simplifying results in

$$p\lambda_{w(S)} + q\lambda_{w(T)} \geq q\lambda_{s(S)} + p\lambda_{s(T)},$$

leading to a contradiction since  $\lambda_{w(S)} < \lambda_{s(T)}$  and  $\lambda_{w(T)} < \lambda_{s(S)}$ . Consequently,  $s_p$  is strictly better off participating in  $T'$  or  $t_q$  is strictly better off forming  $S'$ .

We have shown that  $f_{S'}(y) \neq \emptyset$  or  $f_{T'}(y) \neq \emptyset$ , contradicting that  $y \in C$ . Consequently,  $y$  satisfies coalitional assortativity.

Suppose  $y$  violates exclusion of the strong. If  $|Z(y)| \geq \rho$ , then we have  $f_{Z(y)}(y) \neq \emptyset$ , a contradiction. It remains to consider the case  $|Z(y)| < \rho$ . There exists  $j \in Z(y)$ ,  $S \in \mathcal{P}(y)$ , and  $i \in S$  such that  $\lambda_j < \lambda_i$ . We write  $S = \{s_1, \dots, s_p\}$  with  $s_1 < \dots < s_p$ . Consider the coalition  $T = \{j, s_2, \dots, s_p\}$ . We argue that  $f_T(y) \neq \emptyset$ . Notice that  $m_j(Z(y)) = 0 < m_j(T)$ , so  $j$  has an incentive to deviate. It remains to be shown that also members of  $\{s_2, \dots, s_p\} = S \cap T$  have an incentive to deviate. For all  $i \in S \cap T$ ,  $m_i(S) = m_i(T) > 0$ . Since  $|S| = |T|$ , the following inequality holds,

$$\bar{u}(S) = \frac{\lambda_{s_1} + \sum_{i=2}^p \lambda_{s_i}}{|S|} > \frac{\lambda_j + \sum_{i=2}^p \lambda_{s_i}}{|T|} = \bar{u}(T).$$

Therefore, for all  $i \in S \cap T$ , it holds that  $r_i(S) = m_i(S) - \bar{u}(S) < m_i(T) - \bar{u}(T) = r_i(T)$ . We have obtained a contradiction with  $y \in C$ . Consequently,  $y$  satisfies exclusion of the strong.

### Part 3. The Myopic Stable Set coincides with the Core.

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

The proof of iterated external stability is constructive using the following procedure, which, given an initial state  $x_0 \in X \setminus C$ , generates a finite sequence of productive coalitions  $S^1, \dots, S^\ell$  and states  $x_1, \dots, x_\ell$  such that, for  $k = 1, \dots, \ell$ ,  $x_k \in \tilde{f}_{S^k}(x_{k-1})$  and  $x_\ell$  is a bottom-up segregated state. The coalitions  $S^k$  are all disjoint and chosen such that their members have the lowest productivity among players in  $N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa)$ .

Start with a state  $x_0 \in X \setminus C$ . Parts (i) and (ii) formalize a sequence of dominations where first a productive coalition with weak individuals forms and so forth until a strongest productive coalition forms. We define  $k = 1$ .

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

(ii) Consider a coalition  $T \subseteq N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa$  with  $|T| = \rho$  such that for all  $i \in T$  and for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup T)$  it holds that  $\lambda_i \leq \lambda_j$ . If  $f_T(x_{k-1}) \neq \emptyset$ , then define  $S^k = T$ . If  $f_T(x_{k-1}) = \emptyset$ , then there is  $i \in T$  who belongs to a productive coalition  $S \in x_{k-1}$  such that  $r_i(S) \geq r_i(T)$ . For all  $i \in S$ , for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup S)$ , it holds that  $\lambda_i \leq \lambda_j$ . In this case, define  $S^k = S$ . Define  $x_k = y(x_{k-1}, S^k)$ . Increase  $k$  by 1 and move back to (i).

The procedure takes  $\ell \leq n/\rho$  steps and finally generates the state  $x_\ell$ . We have that  $\mathcal{P}(x_\ell) = \{S^k \mid k \in \{1, \dots, \ell\}\}$ . It holds that  $x_\ell \in \tilde{f}^\ell(x_0)$  since, for every  $k \in \{1, \dots, \ell\}$ ,

$x_k \in f_{S^k}(x_{k-1})$  or  $x_k = x_{k-1}$ . It remains to be shown that  $x_\ell$  belongs to the Core. By Parts 1 and 2 of the proof, we need to show that  $x_\ell$  is bottom-up segregated.

We have that  $\mathcal{P}(x_\ell) = \{S^k \mid k \in \{1, \dots, \ell\}\}$ . By construction it holds that  $\lambda_{s(S^1)} \leq \lambda_{w(S^2)} \leq \dots \leq \lambda_{s(S^{\ell-1})} \leq \lambda_{w(S^\ell)}$ . It directly follows that for all  $S, T \in \mathcal{P}(x_\ell)$  with  $S \neq T$ , we have  $\lambda_{w(S)} \geq \lambda_{s(T)}$  or  $\lambda_{w(T)} \geq \lambda_{s(S)}$ , so  $x_\ell$  satisfies coalitional assortativity.

In the procedure,  $x_\ell$  can only contain a coalition of size larger than  $\rho$  if for some  $k = 1, \dots, \ell$ , it holds in Step (ii) of the procedure that  $f_T(x_{k-1}) = \emptyset$ . Since there is  $i \in S^k \cap T$  such that  $r_i(S^k) \geq r_i(T)$ , it holds that  $\bar{u}(S^k) \leq \bar{u}(T)$ . If  $S^k$  contains more than  $\rho$  individuals and there is  $i, j \in S^k$  such that  $\lambda_i \neq \lambda_j$ , then  $\bar{u}(S^k) > \bar{u}(S^k \setminus \{s(S^k)\}) \geq \bar{u}(T)$ , leading to a contradiction. Consequently,  $x_\ell$  satisfies homophily.

If  $Z(x_\ell) \neq \emptyset$ , then  $|Z(x_\ell)| < \rho$ , as otherwise the procedure does not stop. By construction, it holds that for every  $i \in \cup_{k=1}^\ell S^k$ , for every  $j \in N \setminus \cup_{k=1}^\ell S^k = Z(x_\ell)$  that  $\lambda_i \leq \lambda_j$ , so  $x_\ell$  satisfies exclusion of the strong.  $\square$

## Proof of Theorem 65

The following lemma is helpful in the proof of Theorem 65.

**Lemma 70.** *Let  $\Gamma = (N, X, E, \{\geq_i\}_{i \in N})$  be a social environment induced by a competitive society  $(N, \lambda, \rho)$ . Let  $x \in X$  and  $S \in x$  be such that  $|S| > \rho$  and  $\lambda_i \neq \lambda_j$  for some  $i, j \in S$ . Then  $f_{S \setminus \{w(S)\}}(x) \neq \emptyset$ .*

*Proof.* Consider the coalition  $T = S \setminus \{w(S)\}$ . Since  $|T| \geq \rho$ , in order to prove that  $f_{S \setminus \{w(S)\}} \neq \emptyset$ , it suffices to show that  $m_i(T) > m_i(S)$  for all  $i \in T$ . Observe that  $\lambda_{w(S)} < \sum_{i \in S} \lambda_i / |S|$  and thus  $\bar{u}(T) > \bar{u}(S)$ . It follows that, for every  $i \in T$ ,  $m_i(T) = \bar{u}(T) > \bar{u}(S) = m_i(S)$ .  $\square$

**Proof of Theorem 65.** The proof consists of three parts. We show first that  $\bar{X} \subseteq C$ , next that  $\bar{X} \supseteq C$ , thereby showing that the Core coincides with the set of top-down segregated states, and finally that the Myopic Stable Set coincides with the Core.

**Part 1.**  $\bar{X} \subseteq C$ .

Let  $\bar{x} \in X$  be a top-down segregated state, where  $\bar{x} = \{S^1, \dots, S^m\}$ ,  $\mathcal{P}(\bar{x}) = \{S^1, \dots, S^\ell\}$ , and, for every  $k \in \{1, \dots, \ell - 1\}$ ,  $\lambda_{s(S^k)} \leq \lambda_{w(S^{k+1})}$ . Moreover, we write  $S^1 = \{s_1^1, \dots, s_p^1\}$  with  $s_1^1 < \dots < s_p^1$ .

Towards a contradiction, suppose that  $\bar{x} \notin C$ . Thus, there exists a coalition  $T = \{t_1, \dots, t_q\}$  with  $t_1 < \dots < t_q$  and a state  $y$  such that  $y \in f_T(\bar{x})$ . Notice that any deviating coalition has to be productive, i.e.,  $|T| \geq \rho$ . We proceed by induction.

**Step 1:**  $S^1 \cap T = \emptyset$ .

Notice that coalition  $S^1$  is productive. We consider two cases: a case in which, for all  $i, j \in S^1$ ,  $\lambda_i = \lambda_j$  and a case in which there are some  $i, j \in S^1$  with  $\lambda_i \neq \lambda_j$ .

In the former case, we have that  $\bar{u}(S^1) = \lambda_1 \geq \bar{u}(T)$  as, by assumption, 1 is a strong individual in the society. Therefore, it holds that  $S^1 \cap T = \emptyset$ .

In the latter case, homophily of  $\bar{x}$  implies that  $|S^1| = \rho$ . We define the collection of  $\rho$  strongest individuals in  $T$  by  $T' = \{t_1, \dots, t_\rho\}$ . It holds that  $\bar{u}(T') \geq \bar{u}(T)$ . For  $i = 1, \dots, \rho$ , it holds that  $\lambda_{s_i} \geq \lambda_{t_i}$ , so we have that  $\bar{u}(S^1) \geq \bar{u}(T') \geq \bar{u}(T)$ , and it follows that  $S^1 \cap T = \emptyset$ .

**Step 2:** If  $\bigcup_{\kappa=1}^k S^\kappa \cap T = \emptyset$ , then  $S^{k+1} \cap T = \emptyset$ .

If  $k \geq \ell$ , then  $S^{k+1} \notin \mathcal{P}(\bar{x})$ . As  $\bigcup_{\kappa=1}^k S^\kappa \cap T = \emptyset$  and  $|Z(\bar{x})| < \rho$  it follows that  $|T| < \rho$ , a contradiction to  $|T| \geq \rho$ .

If  $k < \ell$ , then  $S^{k+1} \in \mathcal{P}(\bar{x})$  and we write  $S^{k+1} = \{s_1^{k+1}, \dots, s_p^{k+1}\}$  with  $s_1^{k+1} < \dots < s_p^{k+1}$ . If for all  $i, j \in S^{k+1}$ ,  $\lambda_i = \lambda_j$ , then  $\bar{u}(S^{k+1}) = \lambda_{s_1^{k+1}} \geq \bar{u}(T)$  as, by assumption,  $s_1^{k+1}$  is a strong individual in  $N \setminus \bigcup_{\kappa=1}^k S^\kappa$ . Therefore, for all  $i \in S^{k+1} \cap T$ , it holds that  $m_i(S^{k+1}) \geq m_i(T)$ , so  $S^{k+1} \cap T = \emptyset$ . If there is some  $i, j \in S^{k+1}$  with  $\lambda_i \neq \lambda_j$ , by homophily,  $|S^{k+1}| = \rho$ . We define  $T' = \{t_1, \dots, t_\rho\}$ , so  $\bar{u}(T') \geq \bar{u}(T)$ . Since  $\lambda_{s_i^{k+1}} \geq \lambda_{t_i}$  for  $i = 1, \dots, \rho$ , we have that  $\bar{u}(S^{k+1}) \geq \bar{u}(T') \geq \bar{u}(T)$ . It follows that  $S^{k+1} \cap T = \emptyset$ .

By Steps 1 and 2 it follows that  $T = \emptyset$  and we have obtained a contradiction. Consequently, it holds that  $\bar{x} \in C$ .

**Part 2.**  $\bar{X} \supseteq C$ . Let some  $y \in C$  be given. Towards a contradiction, suppose that  $y$  is not top-down segregated.

If homophily is violated, then there exists a coalition  $S \in y$  with  $|S| > \rho$  such that for some  $i, j \in S$ ,  $\lambda_i > \lambda_j$ . Thus, by Lemma 70, it holds that  $f_{S \setminus \{w(S)\}}(y) \neq \emptyset$ , a contradiction to  $y$  being in the Core.

If coalitional assortativity is violated, then there are coalitions  $S, T \in \mathcal{P}(y)$  such that  $\lambda_{w(S)} < \lambda_{s(T)}$  and  $\lambda_{w(T)} < \lambda_{s(S)}$ . We write  $S = \{s_1, \dots, s_p\}$  and  $T = \{t_1, \dots, t_q\}$  with  $s_1 < \dots < s_p$  and  $t_1 < \dots < t_q$ . We show that for  $S' = \{s_1, \dots, s_{p-1}, t_1\}$  it holds that  $f_{S'}(y) \neq \emptyset$  or for  $T' = \{t_1, \dots, t_{q-1}, s_1\}$  it holds that  $f_{T'}(y) \neq \emptyset$ .

Since  $|S| = |S'| \geq \rho$  and  $\lambda_{q_1} > \lambda_{s_p}$ , we have, for all  $i \in S \cap S' = S' \setminus \{t_1\}$ ,

$$m_i(S) = \bar{u}(S) = \frac{\lambda_{s_p} + \sum_{i=1}^{p-1} \lambda_{s_i}}{|S|} < \frac{\lambda_{y_1} + \sum_{i=1}^{p-1} \lambda_{s_i}}{|S'|} = \bar{u}(S') = m_i(S').$$

By symmetry, it also holds for all  $i \in T \cap T' = T' \setminus \{s_1\}$  that  $m_i(T) < m_i(T')$ .

Towards a contradiction, suppose that  $s_1 = s(S)$  has no incentive to participate in  $T'$  and  $t_1 = s(T)$  has no incentive to form  $S'$ , so  $\bar{u}(S) \geq \bar{u}(T')$  and  $\bar{u}(T) \geq \bar{u}(S')$ . Thus, by definition,

$$\begin{aligned} \frac{\sum_{i \in S} \lambda_i}{p} &\geq \frac{\lambda_{s(S)}}{q} + \frac{\sum_{i \in T \setminus \{w(T)\}} \lambda_i}{q}, \\ \frac{\sum_{i \in T} \lambda_i}{q} &\geq \frac{\lambda_{s(T)}}{p} + \frac{\sum_{i \in S \setminus \{w(S)\}} \lambda_i}{p}. \end{aligned}$$

These expressions can be simplified to get

$$\begin{aligned} q \sum_{i \in S} \lambda_i &\geq p \lambda_{s(S)} + p \sum_{i \in T \setminus \{w(T)\}} \lambda_i, \\ p \sum_{i \in T} \lambda_i &\geq q \lambda_{s(T)} + q \sum_{i \in S \setminus \{w(S)\}} \lambda_i. \end{aligned}$$

Adding up these two inequalities and simplifying results in

$$q \lambda_{w(S)} + p \lambda_{w(T)} \geq p \lambda_{s(S)} + q \lambda_{s(T)},$$

leading to a contradiction since  $\lambda_{w(S)} < \lambda_{s(T)}$  and  $\lambda_{w(T)} < \lambda_{s(S)}$ . Consequently,  $s_1$  is strictly better off participating in  $T'$  or  $t_1$  is strictly better off forming  $S'$ .

We have shown that  $f_{S'}(y) \neq \emptyset$  or  $f_{T'}(y) \neq \emptyset$ , contradicting that  $y \in C$ . Consequently,  $y$  satisfies coalitional assortativity.

Suppose  $y$  violates exclusion of the weak. If  $|Z(y)| \geq \rho$ , then we have  $f_{Z(y)}(y) \neq \emptyset$ , a contradiction. It remains to consider the case  $|Z(y)| < \rho$ . There exists  $j \in Z(y)$ ,  $S \in \mathcal{P}(y)$ , and  $i \in S$  such that  $\lambda_i < \lambda_j$ . We write  $S = \{s_1, \dots, s_p\}$  with  $s_1 < \dots < s_p$ . Consider the coalition  $T = \{s_1, \dots, s_{p-1}, j\}$ . We argue that  $f_T(y) \neq \emptyset$ . Notice that  $m_j(Z(y)) = 0 < m_j(T)$ , so  $j$  has an incentive to deviate. It remains to be shown that also members of  $\{s_1, \dots, s_{p-1}\} = S \cap T$  have an incentive to deviate. Since  $|S| = |T|$ , for all  $i \in S \cap T$  the following inequality holds,

$$m_i(S) = \bar{u}(S) = \frac{\lambda_{s_p} + \sum_{i=1}^{p-1} \lambda_i}{|S|} < \frac{\lambda_j + \sum_{i=1}^{p-1} \lambda_i}{|T|} = \bar{u}(T) = m_i(T).$$

We have obtained a contradiction with  $y \in C$ . Consequently,  $y$  satisfies exclusion of the weak.

### Part 3. The Myopic Stable Set coincides with the Core.

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

The proof of iterated external stability is constructive using the following procedure, which, given an initial state  $x_0 \in X \setminus C$ , generates a finite sequence of productive coalitions  $S^1, \dots, S^\ell$  and states  $x_1, \dots, x_\ell$  such that, for  $k = 1, \dots, \ell$ ,  $x_k \in \tilde{f}_{S^k}(x_{k-1})$  and  $x_\ell$  is a top-down segregated state. The coalitions  $S^k$  are all disjoint and chosen such that their members have the highest productivity among players in  $N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa)$ .

Start with a state  $x_0 \in X \setminus C$ . We define  $k = 1$ .

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

(ii) Consider a coalition  $T \subseteq N \setminus \bigcup_{\kappa=1}^{k-1} S^\kappa$  with  $|T| = \rho$  such that for all  $i \in T$  and for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup T)$  it holds that  $\lambda_i \geq \lambda_j$ . If  $f_T(x_{k-1}) \neq \emptyset$ , then define  $S^k = T$ . If  $f_T(x_{k-1}) = \emptyset$ , then there is  $i \in T$  who belongs to a productive coalition  $S \in x_{k-1}$  such that  $m_i(S) \geq m_i(T)$ . For all  $i \in S$ , for all  $j \in N \setminus (\bigcup_{\kappa=1}^{k-1} S^\kappa \cup S)$ , it holds that  $\lambda_i \geq \lambda_j$ . In this case, define  $S^k = S$ . Define  $x_k = y(x_{k-1}, S^k)$ . Increase  $k$  by 1 and move back to (i).

The iteration takes  $\ell \leq n/\rho$  steps and finally generates the state  $x_\ell$ . We have that  $\mathcal{P}(x_\ell) = \{S^k \mid k \in \{1, \dots, \ell\}\}$ . It holds that  $x_\ell \in \tilde{f}^\ell(x_0)$  since, for every  $k \in \{1, \dots, \ell\}$ ,  $x_k \in f_{S^k}(x_{k-1})$  or  $x_k = x_{k-1}$ . It remains to be shown that  $x_\ell$  belongs to the Core. By Parts 1 and 2 of the proof, we need to show that  $x_\ell$  is top-down segregated.

We have that  $\mathcal{P}(x_\ell) = \{S^k \mid k \in \{1, \dots, \ell\}\}$ . By construction it holds that  $\lambda_{w(S^1)} \geq \lambda_{s(S^2)} \geq \dots \geq \lambda_{w(S^{\ell-1})} \geq \lambda_{s(S^\ell)}$ . It directly follows that for all  $S, T \in \mathcal{P}(x_\ell)$  with  $S \neq T$ , we have  $\lambda_{w(S)} \geq \lambda_{s(T)}$  or  $\lambda_{w(T)} \geq \lambda_{s(S)}$ , so  $x_\ell$  satisfies coalitional assortativity.

In the procedure,  $x_\ell$  can only contain a coalition of size larger than  $\rho$  if for some  $k = 1, \dots, \ell$ , it holds in Step (ii) of the procedure that  $f_T(x_{k-1}) = \emptyset$ . Since there is  $i \in S^k \cap T$  such that  $m_i(S^k) \geq m_i(T)$ , it holds that  $\bar{u}(S^k) \geq \bar{u}(T)$ . If  $S^k$  contains more than  $\rho$  individuals and there is  $i, j \in S^k$  such that  $\lambda_i \neq \lambda_j$ , then  $\bar{u}(S^k) < \bar{u}(S^k \setminus \{w(S^k)\}) \leq \bar{u}(T)$ , leading to a contradiction. Consequently,  $x_\ell$  satisfies homophily.

If  $Z(x_\ell) \neq \emptyset$ , then  $|Z(x_\ell)| < \rho$ , as otherwise the procedure does not stop. By construction, it holds that for every  $i \in \cup_{k=1}^{\ell} S^k$ , for every  $j \in N \setminus \cup_{k=1}^{\ell} S^k = Z(x_\ell)$  that  $\lambda_i \geq \lambda_j$ , so  $x_\ell$  satisfies exclusion of the weak.  $\square$

# References

- ACEMOGLU, D., G. EGOROV, AND K. SONIN (2008), "Coalition Formation in Non-Democracies", *Review of Economic Studies*, 75, 987–1009.
- ALIPRANTIS, C.D., AND K.C. BORDER (2006), *Infinite Dimensional Analysis - A Hitchhiker's Guide*, Springer, Berlin.
- ASHLAGI, I., Y. KANORIA, AND J.D. LESHNO (2017), "Unbalanced Random Matching Markets: The Stark Effect of Competition" *Journal of Political Economy*, 125, 69–98.
- AUMANN, R.J. (1959), "Acceptable Points in General Cooperative  $n$ -Person Games," in A.W. Tucker and R.D. Luce (eds.), *Contributions to the Theory of Games, Volume IV*, Princeton University Press, Princeton, pp. 287–324.
- BALA, V., AND S. GOYAL (2000), "A Noncooperative Model of Network Formation," *Econometrica*, 68, 1181–1229.
- BÉAL, S., E. RÉMILA, AND P. SOLAL (2013), "Accessibility and Stability of the Coalition Structure Core," *Mathematical Methods of Operations Research*, 78, 187–202.
- BANERJEE, S., H. KONISHI, AND T. SÖNMEZ (2001), "Core in a Simple Coalition Formation Game," *Social Choice and Welfare*, 18, 135–153.
- BARBERÀ, S., C. BEVIÁ, AND C. PONSATÍ (2015), "Meritocracy, Egalitarianism and the Stability of Majoritarian Organizations," *Games and Economic Behavior*, 91, 237–257.
- BERTRAND, J. (1883), "Review of 'Théorie Mathématique de la Richesse Sociale' and 'Recherches sur les Principes Mathématiques de la Théorie des Richesses'", *Journal des Savants*, 67, 499–508.
- BEVIÁ, C., L. CÓRCHON, AND A. ROMERO-MEDINA (2017), "Relinquishing Power, Exploitation and Political Unemployment in Democratic Organizations," *Social Choice and Welfare*, 49, 735–753.
- BLUME, A. (2003), "Bertrand without Fudge," *Economics Letters*, 78, 167–168.
- BOGOMOLNAIA, A., AND M.O. JACKSON (2002), "The Stability of Hedonic Coalition Structures," *Games and Economic Behavior*, 38, 201–230.
- BONDAREVA, O.N. (1963), "Some Applications of Linear Programming Methods to the Theory of Cooperative Games," *Problemy Kibernetiki*, 10, 119–139.



- BRACHA, A., U. GNEEZY, AND G. LOEWENSTEIN (2015), "Relative Pay and Labor Supply," *Journal of Labor Economics*, 33, 297–315.
- CARD, D., A. MAS, E. MORETTI, AND E. SAEZ (2012), "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review*, 102, 2981–3003.
- CHWE (1994), "Farsighted Coalitional Stability," *Journal of Economic Theory*, 63, 299–325.
- CLARK, A.E., P. FRIJTERS, AND M.A. SHIELDS (2008), "Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles," *Journal of Economic Literature*, 46, 95–144.
- DEEMEN, A.M.A. VAN (1991), "A Note on Generalized Stable Sets," *Social Choice and Welfare*, 8, 255–260.
- DEMUYNCK, T., P.J.J. HERINGS, R.D. SAULLE, AND C. SEEL (2019), "The Myopic Stable Set for Social Environments," *Econometrica*, 87, 111–138
- DINDOŠ, M., AND C. MEZZETTI (2006), "Better-reply Dynamics and Global Convergence to Nash Equilibrium in Aggregative Games," *Games and Economic Behavior*, 54, 261–292.
- DUBEY, P., O. HAIMANKO, AND A. ZAPECHELNYUK (2006), "Strategic complements and substitutes, and potential games," *Games and Economic Behavior*, 54, 77–94.
- DUESENBERY, J.S. (1949), *Income, Saving, and the Theory of Consumer Behavior*, Harvard University Press, Cambridge.
- DUFWENBERG, M., P. HEIDHUES, G. KIRCHSTEIGER, F. RIEDEL, AND J. SOBEL (2011), "Other-regarding Preferences in General Equilibrium," *Review of Economic Studies*, 78, 613–639.
- DUTTA, B., S. GHOSAL, AND D. RAY (2005), "Farsighted Network Formation," *Journal of Economic Theory*, 122, 143–164.
- DUTTA, B., AND S. MUTUSWAMI (1997), "Stable Networks," *Journal of Economic Theory*, 76, 322–344.
- DUTTA, B., AND R. VOHRA (2017), "Rational Expectations and Farsighted Stability," *Theoretical Economics*, 12, 1191–1227.
- FERRER-I-CARBONELL, A. (2005), "Income and Well-being: An Empirical Analysis of the Comparison Income Effect," *Journal of Public Economics*, 89, 997–1019.
- FRANK, R.H. (1985), *Choosing the Right Pond: Human Behavior and the Quest for Status*, Oxford University Press, New York.
- FRIEDMAN, J.W., AND C. MEZZETTI (2001), "Learning in Games by Random Sampling," *Journal of Economic Theory*, 98, 55–84.
- GALE, D., AND L.S. SHAPLEY (1962), "College Admissions and the Stability of Marriage," *American Mathematical Monthly*, 69, 9–15.

- GOYAL, S., P. HERNÁNDEZ, G. MARTÍNEZ-CÁNOVAS, F. MOISAN, M. MUÑOZ-HERRERA, AND A. SÁNCHEZ (2018), "Integration and Diversity," Working Paper, 1–56.
- GREENBERG, J. (1990), "The Theory of Social Situations: An Alternative Game-Theoretic Approach," *Cambridge University Press, Cambridge*.
- HAFER, C. (2006), "On the Origins of Property Rights: Conflict and Production in the State of Nature," *Review of Economic Studies*, 73, 119–143.
- HARSANYI, J.C. (1974), "An Equilibrium-point Interpretation of Stable Sets and a Proposed Alternative Definition," *Management Science*, 20, 1472–1495.
- HART, S., AND M. KURZ (1983), "Endogenous Formation of Coalitions," *Econometrica*, 51, 1047–1064.
- HERINGS, P.J.J., A. MAULEON, AND V. VANNETELBOSCH (2004), "Rationalizability for Social Environments," *Games and Economic Behavior*, 49 135–156.
- HERINGS, P.J.J., A. MAULEON, AND V. VANNETELBOSCH (2009), "Farsightedly Stable Networks," *Games and Economic Behavior*, 67, 526–541.
- HERINGS, P.J.J., A. MAULEON, AND V. VANNETELBOSCH (2017), "Stable Sets in Matching Problems with Coalitional Sovereignty and Path Dominance," *Journal of Mathematical Economics*, 71, 14–19.
- HERINGS, P.J.J., A. MAULEON, AND V. VANNETELBOSCH (2018), "Stability of Networks under Horizon- $K$  Farsightedness," forthcoming in *Economic Theory*.
- JACKSON, M.O., AND A. VAN DEN NOUWELAND (2005), "Strongly Stable Networks," *Games and Economic Behavior*, 51, 420–444.
- JACKSON, M.O., AND A. WATTS (2002), "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, 106, 265–295.
- JACKSON, M.O., AND A. WOLINSKY (1996), "A Strategic Model of Social and Economic Networks," *Journal of Economic Theory*, 71, 44–74.
- JENSEN, M.K. (2010), "Aggregative Games and Best-reply Potentials," *Economic Theory*, 43, 45–66.
- KARTIK, N. (2011), "A Note on Undominated Bertrand Equilibria," *Economics Letters*, 111, 125–126.
- KNUTH, D.E. (1976), *Marriages Stables*, Les Presses de l'Université de Montreal, Montreal.
- KÓCZY, L.Á., AND L. LAUWERS (2004), "The Coalition Structure Core Is Accessible," *Games and Economic Behavior*, 48, 86–93.
- KONISHI, H., AND D. RAY (2003), "Coalition Formation as a Dynamic Process," *Journal of Economic Theory*, 110, 1–41.

- LUCAS, W.F. (1968), "A Game with no Solution," *Bulletin of the American Mathematical Society*, 74, 237–239.
- LUCAS, W.F. (1992), "Von Neumann-Morgenstern Stable Sets," in R.J. Aumann and S. Hart (eds.), *Handbook of Game Theory, Volume 1*, Elsevier Science Publishers, Amsterdam, pp. 543–590.
- LUTTMER, E.F.P. (2005), "Neighbors as Negatives: Relative Earnings and Well-being," *Quarterly Journal of Economics*, 120, 963–1002.
- MACCHERONI, F., M. MARINACCI, AND A. RUSTICHINI (2012), "Social Decision Theory: Choosing within and between Groups," *Review of Economic Studies*, 79, 1591–1636.
- MEYN, S.P., AND R.L. TWEEDIE (1993), *Markov Chains and Stochastic Stability*, Springer-Verlag.
- MILCHTAICH, I., AND E. WINTER (2002), "Stability and Segregation in Group Formation," *Games and Economic Behavior*, 38, 318–346.
- MILGROM, P., AND J. ROBERTS (1990), "Rationalizability, Learning and Equilibrium in Games with Strategic Complementarities," *Econometrica*, 58, 1255–1277.
- MONDERER, D., AND L.S. SHAPLEY (1996), "Potential Games," *Games and Economic Behavior*, 14, 124–143.
- MYERSON, R.B., AND J. WEIBULL (2015), "Tenable Strategy Blocks and Settled Equilibria," *Econometrica*, 83, 943–976.
- MORELLI, M., AND I.-U. PARK (2016), "Internal Hierarchy and Stable Coalition Structures," *Games and Economic Behavior*, 96, 90–96.
- MUTHOO, A. (2004), "A Model of the Origins of Basic Property Rights," *Games and Economic Behavior*, 49, 288–312.
- NEUMANN, J. VON AND O. MORGENSTERN, (1944), "Theory of Games and Economic Behavior," Princeton University Press, Princeton, NJ.
- PAGE, F.H., JR., AND M. WOODERS (2009), "Strategic Basins of Attraction, the Path Dominance Core, and Network Formation Games," *Games and Economic Behavior*, 66, 462–487.
- PAGE, F.H., JR., M. WOODERS, AND S. KAMAT (2005), "Networks and Farsighted Stability," *Journal of Economic Theory*, 120, 257–269.
- PICCIONE, M., AND R. RAZIN (2009), "Coalition Formation under Power Relations," *Theoretical Economics*, 4, 1–15.
- PICCIONE, M., AND A. RUBINSTEIN (2007), "Equilibrium in the Jungle," *Economic Journal*, 117, 883–896.
- RAY, D., AND R. VOHRA (2014), "Coalition Formation," in H.P. Young and S. Zamir (eds.), *Handbook of Game Theory, Volume 4*, Elsevier, Amsterdam, pp. 239–326.
- RAY, D., AND R. VOHRA (2015), "The Farsighted Stable Set," *Econometrica*, 83, 977–1011.

- ROTH, A.E., AND A. POSTLEWAITE (1977), "Weak Versus Strong Domination in a Market with Indivisible Goods", *Journal of Mathematical Economics*, 4, 131–137
- ROTH, A.E., AND J.H. VANDE VATE (1990), "Random Paths to Stability in Two-sided Matching," *Econometrica*, 58, 1475–1480.
- ROTH, A.E., AND E. PERANSON (1999), "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design," *American Economic Review*, 89, 748–780.
- SAWA, R. (2014), "Coalitional Stochastic Stability in Games, Networks and Markets," *Games and Economic Behaviour*, 88, 90–111.
- SCARF, H. (1967), "The Core of an  $N$  Person Game," *Econometrica*, 35, 50–69.
- SELTEN, R. (1970), *Preispolitik der Mehrproduktenunternehmung in der statischen Theorie*, Springer-Verlag, Berlin.
- SHAPLEY, L.S. (1967), "On Balanced Sets and Cores," *Naval Research Logistics Quarterly*, 14, 453–460.
- SHAPLEY, L.S. AND H. SCARF (1974), "On Cores and Indivisibility," *Journal of Mathematical Economics* 1, 23–37.
- SHUBIK, M. (1962), *Game Theory in the Social Sciences: Concepts and Solutions*, MIT Press, Cambridge, Massachusetts.
- TAMURA, A. (1993), "Transformation from Arbitrary Matchings to Stable Matchings," *Journal of Combinatorial Theory, Series A*, 62, 310–323.
- TOPKIS, D. (1979), "Equilibrium Points in Non-Zero Sum  $n$ -Person Submodular Games," *SIAM Journal of Control and Optimization*, 17, 773–787.
- VOORNEVELD, M. (2000), "Best-response Potential Games," *Economics Letters*, 66, 289–295.
- WATTS, A. (2007), "Formation of Segregated and Integrated Groups," *International Journal of Game Theory*, 35, 505–519.
- XUE, L. (1998), "Coalitional Stability under Perfect Foresight," *Economic Theory*, 11, 603–627.



# Valorization Addendum

This addendum discusses the *knowledge valorization* generated by the chapters of the dissertation.

According to the report of the National Valorization Committee,

"knowledge valorization means the process of creating value from knowledge, by making knowledge suitable and/or available for social (and/or economic) use and by making it suitable for translation into competing products, services, processes and new activities."

Examples of social (and/or economic) value creation from research results can be very different from each other depending on the disciplines: licences, open source tools, software, co-publications with social and/or economic stakeholders, publications in journals and newspapers, non-academic publications, appearances in the media, contributions to public debates, advice for social organisations or companies, policy advice for governments and training programmes for professionals.

Of the item listed above, the only one that is directly applicable to the present work is "publications in journals". The paper based upon Chapter 2 was published at *Econometrica* in 2019. The paper based upon Chapter 3 was published at *Theory and Decision* in 2019. The paper based upon Chapter 4 has been submitted to an A journal. However, given the theoretical nature of the dissertation, other opportunities of valorization are merely speculative. Theory is, by its nature, abstract and provides a selective and one-sided account of the many-sided concrete social world.

Theory and research are strongly interrelated. For example, theory enables us to connect a single study to the immense base of knowledge to which other researchers contribute. The framework and results contained in Chapter 2 are exactly in this spirit: at first glance, indeed, literature on coalition formation theory seems particularly fragmented: coalition formation theory belongs to cooperative game theory as well as to non-cooperative game theory. Furthermore, a coalition formation process can be studied by exploiting a variety of settings and solution concepts. In Chapter 2, we provide a new tool which unifies and generalizes different previous results in this area of research. The most interesting feature of this new tool concerns its capability of providing a prediction even in those scenarios where standard tools fail. An example is given by Chapter 3 which studies a Bertrand duopoly with different costs. It is well known that if prices can be any real number a solution in pure strategies does not exist. Despite the popularity of Bertrand duopoly, the set of mixed-strategies Nash equilibria has not been fully characterized, but, infinitely many equilibria exist. We obtain a unique prediction both when predatory pricing is allowed and when it is forbidden by law. An implication of our results is that when

predatory pricing is forbidden, the market price is predicted to be weakly lower than in mixed strategies Nash equilibria.

Chapter 4 is in line with a rising literature in coalition formation theory which studies how social norm and behavioural preferences can affect democratic institutions. In particular we study coalition formation among heterogeneous individuals who care not only about their absolute economic gain but also about their relative economic position within the group they belong to. We consider competitive societies in which the surplus of a coalition is split according to productivity and egalitarian societies in which coalitions split their surplus equally. Both competitive and egalitarian societies lead to segregated partition structures. For competitive societies, all stable allocations are based on bottom-up segregation, i.e., individuals with adjacent productivities form coalitions and if some individuals are not part of a productive coalition, then these are the most productive ones. For egalitarian societies, we obtain top-down segregation in all stable allocations. Again it holds that individuals with adjacent productivities form coalitions, but now the least productive individuals may not be part of any productive coalition. In both societies, if an allocation is not segregated then there exists a sequence of coalitional deviations which lead to a segregated allocation.

However, despite all my effort to extract a social or economic value from these results, I cannot really foresee the direct and indirect implications for the society. In my defence, I would like to conclude this *valorization addendum* borrowing the words of John von Neumann, one of the greatest scientists of all times:

*"A large part of mathematics which becomes useful developed with absolutely no desire to be useful, and in a situation where nobody could possibly know in what area it would become useful; and there were no general indications that it ever would be so."*

# Curriculum Vitae

Riccardo Domenico Saulle was born in Rome on October 31, 1980. He received a B.A.(2009) in Philosophy at La Sapienza University of Rome. He obtained a MBA (2010) in Corporate Social Responsibility at Lumsa University and a two years master degree (2013) in Environmental and Development Economics at Roma Tre University. In 2014, he studied quantitative economics at Luiss University. In 2015, he spent one semester in Maastricht attending advanced courses in Game Theory, Epistemic Game Theory, Public Economics and Behavioural Economics. Between 2015 and 2019 he was PhD candidate at the Department of Economics under the supervision of Jean-Jacques Herings, Christian Seel and Thomas Demuynck. The results of his research are collected in this dissertation. He presented his work in several international conferences. His job market paper is published at *Econometrica* and it has been awarded a Best Student Paper Prize during the 13th European Meeting of Game Theory (SING13). He is post-doctoral researcher at La Sapienza University and he will further is academic career as a post doctoral researcher at Padua University under the supervision of Antonio Nicoló.



He is also graduated in Music Theory (2005) at Santa Cecilia Music Conservatory of Rome and, despite the advise of his lovely wife, he is still persisting to endeavour playing trumpet.

His Erdős number is 3.