**Title**
Multiple-group games



**Author**
Neary, Philip R.

**Publication Date**
2011

UNIVERSITY OF CALIFORNIA, SAN DIEGO

# Multiple-Group Games

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Philip R. Neary

Committee in charge:

Professor Joel K. Sobel, Chair
Professor S. Nageeb Ali
Professor Vincent P. Crawford
Professor Mathew D. McCubbins
Professor Ramamohan Paturi

2011

The dissertation of Philip R. Neary is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

________________________________

________________________________

________________________________

________________________________

________________________________

Chair

University of California, San Diego

2011

DEDICATION

To my mum.

Love p

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I suppose that, in the real world, even Peter Pan has to grow up. Whether or not I do now grow up, I know for sure that I could not have made it this far without lots and lots of help.

To be honest, I have been a burden on the entire UCSD economics department. I wish to thank Devaney Kerr, Rafael Acevedo, Mary-Jane Hubbard, Rebecca Franco, and Mike Bacci for all their help with administrative issues over the years, and for finding me extra TA %s when I probably didn't deserve them. I also enjoyed chatting with each of them on a regular basis as their positive outlooks are quite infectious.

I have been fortunate to work with many outstanding professors. Joel Watson and Hal White kindly gave me RA-ships during my first few summers. Dimitris Politis, Ian Abramson, and in particular Ruth Williams were very generous with their time (all the more so seeing as I was not even in their department). Mohan Paturi and Matt McCubbins were fantastic to have on my committee. Nageeb Ali and Vince Crawford read everything I sent them and were always incredibly supportive - and especially so while I was stressed out on the job market. The fact that I got any job was helped hugely by the ever-kind Julie Cullen.

My friends deserve special mention. Many fun times were had with Youjin, HeeSeung, Hiro, Dalia, Danielken, Oana, Tiffany, Lucas, Ayelen, Tess, and Nick. Daniel and Max were the best roommates I could have imagined and I thank them for putting up with much of my nonsense. Jess and MinSeong would probably each have finished a year earlier were they not so unfortunately located in an office so close to mine (though I still think they secretly loved me stopping by). Schooling Augusto and Justin on the golf course was always fun. Alberto and Chulyoung were fantastic job market comrades. Susan O'Brien took care of me one Christmas and I will always love her for it. Recently, (Professor!) Prashant has joined the contingent and his support during the job market was vaandhi to the max.

I had lots of (unconditional) support from home. My cousins, aunts and uncles, and especially my Gran were always eager to see me whenever I was back. The extended Waterloo Road gang of Jim, David, Eileen, and Marie were always there for me. And I am really looking forward to moving closer to the Oxford crew of my Dad, Mairead, Roisin, and Aisling.

Three people deserve special mention.

My advisor, Joel Sobel, took me in when I was a 3rd year student drifting in the currents. He tried to explain cheap talk to me but unfortunately I didn't get it. Then one day he told me to go and read the papers of KMR and Young in the hope that I might find them interesting - I suppose this dissertation is proof that he was right. He always tried to get me to write clearly, though was only partially successful - anything that makes sense in what follows was most likely a sentence corrected by him. While we had our battles, he always had my best interests at heart, and I will forever consider him a dear friend.

Sarada is my favourite person. It is hard to adequately describe how much she has helped me and been there for me. She is quite simply an angel. MOUSEYYYYYYYY!!!!

My mum (who kicked my hungover self out of the car at Dublin airport some 6 years ago) has both emotionally and financially bankrolled the entire San Diego operation, and I would never have finished were it not for her. We skyped almost every week, and I was always touched that she cared how *I* was doing and not how *work* was going. I love her very much and this dissertation is dedicated to her.

Lastly, to my nemesis in the bio-med library who forever thwarted my attempts to build beds out of library furniture: you have won this battle, but the war is not over.

VITA

| | |
|---|---|
| 2004 | B. A. in Mathematics, Trinity College Dublin |
| 2008 | M. S. in Statistics, University of California, San Diego |
| 2011 | Ph. D. in Economics, University of California, San Diego |

ABSTRACT OF THE DISSERTATION

**Multiple-Group Games**

by

Philip R. Neary

Doctor of Philosophy in Economics

University of California, San Diego, 2011

Professor Joel K. Sobel, Chair

Evolutionary game theory is a formal framework which enables one to model how behaviour in large populations might develop. One of the most promising features of evolutionary game theory is that by assuming agents follow simple rules of thumb, equilibria can emerge as the end result of a dynamic process, rather than just as a static, self-enforcing prescription of play. This in turn allows path-dependent predictions to be made on which equilibrium will be reached. The concept of stochastic stability (Foster and Young, 1990), allows a unique equilibrium to be selected if one also accepts that that the above simple rules of thumb are subject to occasional errors.

My dissertation consists of four essays discussing various aspects of evolutionary models, in particular focusing on equilibrium selection in a new class of

games.

Chapter 1 proposes a new class of stage games, Multiple-Group Games (MGG), which extends existing frameworks by allowing more than one type of pairwise interaction. This surprisingly unstudied feature complicates the analysis slightly but is necessary for describing certain settings with large numbers of heterogeneous agents.

Chapter 2 studies a particular MGG with two groups, The Language Game, in which all pairs of players potentially interact. In particular it focuses on stochastically stable equilibria when the rules of thumb that players follow are myopic best response, and the manner in which they err is uniform. It is shown that three properties affect equilibrium selection: group size, group payoffs, and rates of response.

Chapter 3 looks at The Language Game where all pairs of players are not necessarily connected. It is shown that network architecture immediately affects the set of stochastically stable equilibria. This result seems intuitive and should be contrasted with the homogeneous agent models, in which network structure has no effect on equilibrium selection.

Chapter 4, though technically a paper on equilibrium refinements, examines the ways in which players might make mistakes. It places a simple and *reasonable* constraint on the set of ways that players may err: if strategy A performs worse against the current population behaviour than strategy B, then strategy A is less likely to be played.

# Chapter 1

# Multiple-Group Games

## Abstract

This paper introduces a new class of games, Multiple-Group Games (MGGs), that incorporates (a particular kind of) heterogeneity to existing large population models. The game happens on a graph with vertices representing players, and edges representing pairwise *local-interactions*. The population is partitioned into homogeneous *groups*, with players in a given group defined by how they interact with others. Clearly, local-interactions *within-group* must be constrained symmetric, while those *across-group* need not. Utilities are the sum of payoffs earned in all exchanges where the same action must be used in each. When each local-interaction game of a MGG is a potential game (Shapley and Monderer, 1996), the MGG inherits this property with potential equal to the sum of potentials of each local-interaction. The strategy profile that maximizes this potential is stochastically stable (Foster and Young, 1990) under the Ising Model dynamics of Blume (1993). Finally, I discuss some interesting 2-group, 2-action MGGs that can describe situations ranging from segregation, to technology adoption, to social capital accumulation in ways that existing models cannot.

## 1.1 Introduction

While evolutionary models provide a rich framework for studying how patterns of behaviour develop in societies with large numbers of agents, traditionally almost all attention has been on "evolutionary" explanations of equilibrium.[1] That is, the literature has focused primarily on the many different and interesting ways that large populations of boundedly rational players learn/adapt/evolve to an equilibrium, and less so with the games in which such learning/adapting/evolving occurs. To put it another way, detailed analyses are often restricted to simple classes of games - most frequently settings in which populations interact pairwise via a common $2 \times 2$ normal form game.

The purpose of this paper is to highlight what I feel is a limitation of existing evolutionary models, and to extend them to incorporate (a particular kind of) heterogeneity into the game form. In fact there is nothing "evolutionary" about the contribution in that I add nothing to the way players might play games over time.[2] Rather, other than Section 3.6, I focus almost exclusively on the games people play. Specifically, I propose a new class of games, Multiple-Group Games (MMGs), that are interesting even from the perspective of a one shot playing. MGGs build on the pioneering large but finite population evolutionary models of Kandori, Mailath, and Rob (1993), Young (1993), Ellison (1993), Blume (1993), Morris (2000), and others. The difference between MGGs and these existing models is incredibly simple and was alluded to in the opening paragraph. It can be stated concisely as: there can be more than one type of *local-interaction*.

To contextualize things, the following should help. I view any evolutionary model as having two fundamental components. The first of these is the *stage game* which describes the biological/economic/sociological setting of study. The original and canonical stage game is a one-shot, simultaneous-move game, where agents

[1]Evolutionary Game Theory began with the seminal work of Smith and Price (1973), and has now permeated such fields as biology, economics, and sociology. For textbook treatments, see Smith (1982), Weibull (1995), Hofbauer and Sigmund (1998), and more recently Sandholm (2011).

[2]However, one interesting aspect of proposing a new class of stage games is that there is no reason why existing dynamic adjustment rules can not be applied. Although by adding heterogeneity, existing dynamics may need to be modified - and sometimes in interesting ways.

in the population are randomly matched to interact pairwise. It is then supposed that the stage game gets repeated and that the players are boundedly rational. Undiscounted payoffs are received every period, and strategies for tomorrow must be chosen at the end of today. The second component, which fully characterizes the evolutionary model, is to specify the exact *dynamic adjustment rule*(*s*) by which the players choose strategies for the forthcoming period.[3] This generates population dynamics, with the focus being on which equilibrium decentralized play will lead to. The basic tenet common to all evolutionary dynamics stems from behaviour at the level of the individual - strategies that performed better in the past, should be better represented in the future. Evolutionary dynamics differ across many dimensions, often due to the meaning of the word "better" in the previous sentence.[4]

So to summarize: the stage games of evolutionary game theory are typically simple, the beauty of the models and the insights gained from the models, lie in the exact specification of dynamic, adaptive, learning processes by which an equilibrium can emerge in a large population of boundedly rational agents given sufficient time. That is pretty much all I have to say about dynamics in this paper. I will now review how the stage games of evolutionary models have progressed.

The original finite population stage games, due to Kandori, Mailath, and Rob (1993) (hereafter KMR), are one-shot games in which all players from a large but finite *homogeneous* population of identical agents are randomly matched.[5] (These are really just finite population versions of the continuum-population stage games of Smith and Price (1973).) Due to the homogeneity of the population, the pairwise local-interaction is by definition constrained symmetric.[6]

[3]This is what differentiates a repeated setting in an evolutionary framework from the literature on Repeated Games.

[4]A variety of such *dynamics* exist in the literature: players may *best respond* to various intervals of past play; players may look around and *imitate* their neighbour who is currently doing best; players may adopt the action that they themselves have used most often (*reinforcement learning*); etc. See Fudenberg and Levine (1998) and Young (2005).

[5]It is not hard to see that the action that maximizes expected payoff in a random matching model, is identical to that which maximizes payoff in a model where the population resides on a fully connected graph and each individual *plays the field* using the same action with everyone. Since strategic behaviour is identical in both, I adopt the full-connectedness story from here on out.

[6]The famous "evolutionary" result of KMR, extended to more general settings in Young

As a concrete example, suppose the local-interaction is intended to describe a setting in which players choose operating systems (Katz and Shapiro, 1985; Farrell and Saloner, 1985; Arthur, 1989). Such industries have positive network effects - the greater the number of others using the same system, the greater the value use of that system brings. Well, it must then be the case that all players in the population have the same preference over operating systems. This might be realistic for certain markets but hardly seems all encompassing. For example, in the Mac vs PC market, it is almost unfathomable to think that there is unanimous agreement. Furthermore, it is not always appropriate to model the setting as one in which everyone interacts with everyone else.

Clearly then, adding heterogeneity of some sort is needed. Initial attempts at "heterogeneizing" these simple large but finite population stage games proceeded along one of three different avenues. The first, pioneered by Ellison (1993), Blume (1993), and Morris (2000), is the network model. These papers assume that each player occupies a distinct vertex of a graph, and interacts exclusively with his/her *neighbours*, rather than with the population at large. The heterogeneity across agents stems from the characteristics of the vertex at which they reside, but the pairwise local-interaction that occurs along each edge remains common and symmetric. Precisely, agents are differentiated by the *number of* and the *identity of* the neighbours with whom they interact. But note that agents are <u>not</u> differentiated by how they interact. I call such a situation a *Group Game*.[7]

The limitation of Group Games is immediately clear. That the same symmetric local-interaction occurs between every pair of players restricts the population to be in some sense *homogeneous*. Referring back to the example on choice of operating systems where the common local-interaction is one of pure coordination, all players still share preferences over local-interaction equilibrium profiles, and so if they could coordinate their behaviour, it is obvious that uniform adoption of

(1993), is that if players also occasionally experiment, then population behaviour will lock in on, or "select", a subset of the possibly multiple equilibria - the so-called *stochastically stable equilibria* (Foster and Young, 1990). Perhaps most famously, when the local-interaction is a $2 \times 2$ game of pure coordination and players are myopic best-responders who experiment each period with equal likelihood, this selected equilibrium is the risk-dominant one (which need not coincide with the pareto-dominant one).

[7]Note that the model of KMR is also a Group Game.

the locally pareto-dominant equilibrium action is optimal for one and all. Thus, in this case, the issue of which strategy profile players would like to coordinate on remains uninteresting. What is still of interest however, is whether or not decentralized play will converge on such an outcome, and how network structure might affect the convergence.[8]

The second approach adds heterogeneity to preferences. It supposes the population is split into two distinct groups of agents. Each player interacts pairwise exclusively with agents from the other group but not with those from his/her own group. (Note that this constrains the network to be a bipartite graph.) This allows, but does not require, the local-interaction to be asymmetric. I call such games *Finite Asymmetric Contests* (FACs). Asymmetric contests with infinite sized populations and random matching were first studied in Samuelson and Zhang (1992) and Samuelson (1994).

FACs are interesting because they permit tensions in the local-interaction game. This is true even if it is one of coordination. For example, one group could be men, the other group women, and the local-interaction could be Battle of the Sexes. In applications, one group might be buyers with actions {buy, don't buy}, and the other group sellers with actions {produce High quality good, produce Low quality good}. In the operating systems example, this framework permits the different groups to have different preferences. However, FACs are not suitable for strategic situations on arbitrary networks since the underlying graph is constrained to be bipartite. This means, that despite the potential for asymmetry and heterogeneity in the local-interaction, each player still only ever interacts with one other type of agent (men with women and women with men, sellers with buyers and buyers with sellers, those with preferences for Mac with those with preference for PCs). Thus, it is still the case that whatever local-interaction occurs between any pair of players is common.

The third avenue maintains the local-interaction exchanges, but complicates the stage games further. Oechssler (1997), Oechssler (1999), and Ely (2002) are

[8] The risk-dominance selection result of KMR and Young (1993) was extended to some specific networks by Ellison (1993), and to arbitrary networks by Peski (2010). Ellison (1993) did note however that network architecture can dramatically alter rates of convergence.

models with *endogenous* local-interaction, such that players still interact pairwise, but a strategy is now a choice of location and an action. Kim and Sobel (1995) and Matsui (1991) add a round of costless communication, “cheap talk”, before actions are taken, such that a strategy is a choice of message and an action, or a choice of action conditioned on a received message. While this third avenue allows much richness to be added, the models quickly become intractable with the addition of heterogeneous agents (each of these papers still assume a common symmetric local-interaction). I ignore such complex variants, and stick with situations where a strategy corresponds to a simple choice of action.

The definition of a Multiple-Group Game, fuses both Group Games and Finite Asymmetric Contests, retaining the merits of each, removing the limitations of each, and acquiring some new desirable properties. Similar to FACs, it supposes that the population is partitioned into homogeneous groups. Again, players are located on a graph and interact pairwise. Again, utilities are the sum of payoffs gathered in all local-interactions with neighbours, where the same action must be used in each. The key difference is that a player may interact with those from *within* his/her own group, and also with those from *outside* his/her own group. (Note that this is precluded by all the models mentioned above!) Clearly this requires specifying more than one type of local-interaction. Within-group local-interactions are constrained symmetric, while those across-group are not (modulo the requirement that each player has the same action set available to him/her as he/she does in his/her respective within-group local-interactions). That is it. Nothing more.

There are three appealing and realistic features that MGGs allow. The first of these was touched on previously. That is, there exist situations in which an agent’s action must be used with more than one other type of agent. It is not always appropriate that in a population with multiple groups that agents only interact with those in the other group. For example, recall the coordination problem where people are choosing operating systems. Modeling this with Group Games, it is the case that everyone has the same preference over the systems. This seems limiting. Modeling this with FACs, it is the case that even if the groups’ preferences disagree, each group still only interacts with the other one, which again seems far from all-encompassing. Furthermore, in both these cases, an immediate consequence of having network effects is that the only stable outcomes are for all agents in the population to adopt the same system. But in practice, often multiple systems coexist.

The second is that the value an agent places on different local-interaction action profiles may not be opponent-independent. (Clearly this is the case for both Group Games and FACs since each agent only ever interacts with one type player!) It may be that while an agent wants to coordinate with each of his/her neighbours, he/she values coordination with a certain subset more (perhaps his/her own group). The operating systems setting might be one such example. Further still, consider a social setting where an agent wishes to meet his/her friends. The agent must decide to go to one of two bars. While the agent is happier for each additional person who attends the same bar as he/she does, it is very possible that there is a subset of his/her friends who the agent values meeting more. In a more extreme sense, it may be that an agent's optimal action, and not just his/her payoff, varies with the opponent's decision. For example, as in Schelling's segregation models (Schelling, 1969, 1971), it may be the case that an agent wants to coordinate with one group (his/her own race), and yet anti-coordinate with those in the other group (the other race).

The third and final property is that even if all an agent's local-interactions are opponent-independent, it may not be the case that his/her neighbours do also. One could imagine a setting similar to the segregation models of Schelling, in which one group is trying to coordinate with their own but anti-coordinate with the other as before, while the other group is eager to coordinate with one and all.

I believe the main contribution of the paper to be the formal definition of a Multiple-Group Game. With something so simple as its focus, the paper possibly suffers from the deficiency of whether or not a stand-alone definition can be interesting and/or useful by itself. It is my hope that the definition is rich enough and so easily understood that it is in fact both of the above. While the definition by itself seems too general to prove global results for, I hope to convince the reader that there are many case by case examples for which detailed analyses might be interesting.[9]

Finally, I should point out that viewing the contribution of the paper as the definition of a MGG is dangerous given how close it is to some existing work. To my knowledge, the insight of agents interacting with both those in their own group and those in other groups appeared first in Friedman (1998), who looks at a situation with two distinct continuum sized populations where there are also "own-population effects".[10] The *N-Species Frequency-Dependent Interactions* of Cressman, Garay, and Hofbauer (2001) are in some sense the continuum-population version of MGGs, with the additional constraints that across-species interactions must be asymmetric, and matching is uniform (they cannot allow network structure, since networks have no meaning for infinite populations). Their focus is on dynamics, specifically in characterizing the Evolutionary Stable Strategies (Smith and Price, 1973) of a multiple species interaction.[11] Lastly, the definition of a MGG is in fact identical to that of a "polymatrix game" (Janovskaya, 1968; Howson, 1972; Eaves, 1973), as studied by mathematicians and computer scientists.[12] However, the focus in these fields is quite different as they seek answers to questions like whether or not equilibria for such games can be computed in polynomial time. At the risk of repeating myself, my focus is on specific economic situations that the MGG framework permits case by case analysis of. How players might play these games, the analysis of what long run behaviour emerges, and issues of equilibrium computation in such games is left to future research.

The balance of the paper is as follows. The following Section discusses a simple classroom experiment with random matching that extends existing ones to allow heterogeneity in the population. Section 1.3 introduces my notation and illustrates formally how MGGs build on the existing stage game models of KMR, Young (1993), Ellison (1993), Blume (1993), Morris (2000), and Samuelson and

[9] Neary (2010a) and Neary (2010b) analyse a particular two-action, two-group MGG, The Language Game, for the case of random matching and networked societies respectively. Such a thorough analysis could be performed for all the examples in Section 1.7, and for any other examples that the reader can come up with.

[10] Thanks to both Dan Friedman and David Levine for pointing this out.

[11] Thanks to Bill Sandholm for bringing this to my attention.

[12] Thanks to Bernhard von Stengel for informing me of polymatrix games.

Zhang (1992) among others. Section 1.4 shows that if each local-interaction of a MGG is a potential game (Shapley and Monderer, 1996), then the MGG inherits this property. This is useful since, as I show in Section 3.6, under the Ising model dynamics of Blume (1993), the pure strategy profiles which maximize the potential function are precisely the stochastically stable equilibria. Section 1.6 discusses MGGs with two groups and two actions, $2 \times 2$ MGGS, and formalizes notions of when preferences are aligned or not. Section 1.7 demonstrates the usefulness of $2 \times 2$ MGGs via four real world examples, useful for studying such social phenomena as competing standards, segregation, and social capital accumulation. The examples are purely illustrative and meant only to display how broad the MGG framework is; I do not pretend to give a definitive account of any of these issues. Section 1.8 argues that certain terms used in evolutionary game theory have strange interpretations in economic settings, a result of being carried over from evolutionary biologists. Section 1.9 concludes.

## 1.2 Classroom Experiments

A common classroom experiment designed to illustrate the power of evolutionary forces in large populations goes like this. A professor walks in to class and says something like the following: "Let us now play a game. Everyone will privately choose an action, $a$ or $b$, and then I will randomly select two players. Payoffs in dollars for this pair will be given by the following symmetric *local-interaction* game."

| | $a$ | $b$ |
|---|---|---|
| $a$ | $2, 2$ | $0, 0$ |
| $b$ | $0, 0$ | $2, 2$ |

This situation is then repeated for a number of rounds, with the professor always publicly revealing how everybody played the previous round. The point of the exercise is that while the professor has no idea what will happen in the opening round (that is, expecting immediate equilibrium play in a large population coordination problem is unlikely), he/she can predict that after a few rounds, population behaviour will drift towards one of the two coordinated outcomes where everyone adopts the same action. The professor has then demonstrated that even though population behaviour starts at an arbitrary outcome, equilibrium can be the end result of dynamic, so-called "evolutionary", forces.

Another such classroom experiment goes like this. The professor walks in, divides the classroom down the middle, and says something akin to: "Let us now play a game. Those of you on this [left] side of the room are the 'row' players, those of you on this [right] side of the room are the 'column' players. Everyone will privately choose an action and then I will randomly select one player from either side. Payoffs in dollars for this pair will be given by the following asymmetric *local-interaction* game."

| | | right | |
|---|---|---|---|
| | | $a$ | $b$ |
| left | $a$ | $2,1$ | $0,0$ |
| | $b$ | $0,0$ | $1,2$ |

If this situation were also repeated for a number of rounds, with as before the professor always publicly revealing the distribution of play from the previous round, again the professor would be quite confident that after a small number of repetitions, population behaviour will drift towards one of the two coordinated outcomes with uniform adoption.[13]

In the first of these experiments, the population is homogeneous. Thus, the model is only useful for describing situations where a population is, well, homogeneous. In the second, the population is heterogeneous with two distinct groups of agents. However, while allowing heterogeneity in preferences and a clear tension over desired equilibrium profiles, it is still the case that each agent only ever interacts with one homogeneous group (the other group). From the perspective of a classroom experimenter, this means that once you allow heterogeneity in preferences, it is impossible to allow uniform random matching over the whole class. But

[13] Though Professor Joel Sobel says he would not be confident of this.

this raises some immediate issues: Why does adding heterogeneity always come at the cost of uniform random matching? Is uniform random matching never the appropriate model for heterogeneous populations? And if it is sometimes appropriate, what extensions to the above are needed?

The limitation to both these large population games is actually very easily summarized: there is only one common pairwise interaction. That is, whatever interactions do occur, be they symmetric or asymmetric, whether they involve homogeneous or heterogeneous populations, they are all the same!

The following example is one of a large number of simple classroom experiments that illustrate the advantages that Multiple-Group Games can bring to the analysis of large populations.[14] Now, the professor walks in, divides the classroom down the middle, and says: "We will now play a game. Those of you on this [left] side of the room are Group $A$, those of you on this [right] side of the room are Group $B$. Everyone will privately choose an action and then I will randomly select two players. If these two players are from Group $A$, then payoffs in dollars for this pair will be given by $G^{AA}$ below. If the pair are both from Group $B$ then payoffs will be assigned according to $G^{BB}$ below. If the players are from different groups, then payoffs will be given by $G^{AB}$ where the Group $A$ member is the row player, and the Group $B$ member is the column player."

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | 2, 2 | 0, 0 |
| $b$ | 0, 0 | 1, 1 |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | 2, 1 | 0, 0 |
| $b$ | 0, 0 | 1, 2 |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | 1, 1 | 0, 0 |
| $b$ | 0, 0 | 2, 2 |

This new game has a large heterogeneous population, with players potentially interacting with those from their own group and also with those from the other group. Note that in this game players do not care who they are matched with,

[14]This particular game is the "Language Game" of Neary (2010a). I have run this experiment in class with UCSD undergraduates. In fact, if one can understand the classroom experiments described before, it takes only slightly more reasoning power to understand this MGG.

they only care about the actions chosen once matched. While clearly this need not be the case, for the purpose of illustration, and that of a classroom experiment, it makes things easy to follow.

## 1.3 Multiple-Group Games

A finite $N$-person normal form game is a $2N$-tuple $G = (S_1, \ldots, S_N, U_1, \ldots, U_N)$, where for each $i \in \mathcal{N} := \{1, \ldots, N\}$, $S_i$ is a finite nonempty set of pure strategies,[15] and $U_i : \prod_{j=1}^{N} S_j \to \mathbb{R}$ is a *payoff* or *utility* function. Define $\mathbf{S} := \prod_{j=1}^{N} S_j$ with typical element $\mathbf{s} = (s_1, \ldots, s_N)$, and $\mathbf{S}_{-i} := \prod_{j \neq i} S_j$ with typical element $\mathbf{s}_{-i}$. From player $i$'s perspective a pure strategy profile $\mathbf{s} \in \mathbf{S}$ can be viewed as $(s_i, \mathbf{s}_{-i})$. Thus, $(\hat{s}_i, \mathbf{s}_{-i})$ will refer to the profile $(s_1, \ldots, s_{i-1}, \hat{s}_i, s_{i+1}, \ldots, s_N)$, i.e. the profile $\mathbf{s}$, where $\hat{s}_i$ replaces $s_i$.

A fundamental component of my analysis are *symmetric games*. A symmetric game is one where all players possess identical strategy sets and share a common payoff function. The payoff function is common in the sense that the payoff from playing a particular strategy depends only on the strategies chosen and not on the identities of the players taking those strategies. The formal definition is via equality of payoffs with respect to all *permutations* of the player set (Dasgupta and Maskin, 1986).[16,17]

**Definition 1.** An $N$-player game, $G$, is *symmetric* if $S_1 = \cdots = S_N = S$, and for any permutation $\sigma$ we have that, $U_i(s_i, \mathbf{s}_{-i}) = U_{\sigma(i)}(s_{\sigma(i)}, \mathbf{s}_{-\sigma(i)})$.
A symmetric game can be represented by the 3-tuple $(N, S, U)$, where $U$ is the common payoff function.

Before defining our first type of large population game, a *Group Game*, some terminology is needed.

[15] From here on in, the objects of $S_i$ are actions publicly observable to others. As such the words "strategy" and "action" will be used interchangeably throughout.

[16] A *permutation*, $\sigma$, of a set $Z$, is a bijection from $Z$ to itself.

[17] A common verbal definition given for symmetric games is the following: "a symmetric game is a game where the payoffs for playing a particular strategy depend only on the other strategies employed, not on who is playing them." The word *other* in this definition is perhaps contentious. If you read it as I do, then all 2-player games are symmetric. Furthermore, games like the Language Game described in the previous section are as well.

**Definition 2.** A *undirected graph*, $\Gamma$, is a pair $(\mathcal{N}, E)$, where $\mathcal{N}$ denotes the set of vertices, and $E$ the set of edges $E \subset \{(i,j)|i \neq j \in \mathcal{N}\}$ such that $(j,i) \in E$ whenever $(i,j) \in E$. I write $g_{ij}$ for the indicator function of the event $\{(i,j) \in E\}$, and view $E$ as $\{g_{ij}\}_{i,j \in \mathcal{N}}$. If $g_{ij} = 1$, I will say that $i$ and $j$ are *adjacent*, that edge $(i,j)$ has *ends* $i$ and $j$, and that $j$ $(i)$ is a neighbour of $i$ $(j)$.

A graph is said to be *path-connected*, if for any two vertices $i$ and $j$, there exists a collection of vertices, $\{v_0, v_1, \ldots, v_{n-1}, v_n\}$, with $i = v_0$ and $j = v_n$, such that $g_{v_0 v_1} \cdots g_{v_{n-1} v_n} = 1$.

Now we are ready for the definition of a Group Game.

**Definition 3.** A *Group Game* (GG), is defined as the tuple $\Big\{\mathcal{N}, S, G, \Gamma\Big\}$, where

1. $\mathcal{N}$ is the finite set of players. $|\mathcal{N}| = N \geq 2$.

2. $S$ is a finite-action strategy set, common to all the players.

3. $G = (2, S, u)$ is a symmetric 2-player game known as the *local-interaction*.

4. $\Gamma = (\mathcal{N}, E)$ is a path-connected graph.

The above primitives generate a vector of utility functions, $\mathbf{U} = \{U_1, \ldots, U_N\}$, where for any player $i \in \mathcal{N}$, the utility from taking action $s \in S$ is

$$U_i(s, \mathbf{s}_{-i}) := \sum_{j \neq i} g_{ij}\, u(s, s_j)$$

Simply put, a Group Game is a large population interaction, where players interact pairwise via a common *local-interaction* game that is constrained symmetric. Utilities are given by the sum of payoffs earned with neighbours, where the same action must be used with each. This latter feature of Group Games is very important. It rules out the case where a player uses one action in one exchange, and another different action in another exchange.[18]

[18]This is not only for tractability reasons, but is quite realistic in many situations, since there may be constraints on players that prevent them from conditioning their action on whom they interact with. Morris (2000) goes further saying: "The strategic problem becomes interesting when it is assumed that players cannot tailor their behaviour for each neighbour, but must choose a constant action for all neighbours."

When $g_{ij} = 1$ for all $i \neq j$, I say the Group Game is a *Fully-Connected Group Game*. It is easy to see that in this case the Group Game is a symmetric game as $U_i = U_{\sigma(i)} = U$ for any permutation $\sigma$.

Definition 3 is very broad. The uniform random matching games of KMR and Young (1993) can be viewed as Group Games, as can finite population versions of the network models of Ellison (1993), Blume (1993), and Morris (2000). Similarly, some of the games studied in Galeotti, Goyal, Jackson, Vega-Redondo, and Yariv (2009), and the anti-coordination games of Bramoulle and Kranton (2007) and Bramoulle (2007) are encompassed by the definition.

Group Games are useful in that they allow the study of certain strategic situations on (arbitrary) networks. The players are often referred to *identical* because the local-interaction is symmetric. Due to the varying number of neighbours that each player possesses, this terminology is imprecise but not particularly misleading. Furthermore, each player faces a similar situation because of the pairwise nature of the interactions via a common local-interaction.

The second type of large population interactions are *Finite Asymmetric Contests*. I define these after a little more terminology.

**Definition 4.** A *partition* of $\mathcal{N}$ is a nonempty and pairwise disjoint collection of subsets $\{\pi_1, \ldots, \pi_M\}$ such that $\cup_{l=1}^{M} \pi_l = \mathcal{N}$. I assume the partition elements $\pi_1, \ldots, \pi_M$ are each nonempty and I refer to them as *groups*.

**Definition 5.** A graph, $\Gamma = (\mathcal{N}, E)$, is said to be *bipartite*, if the vertex set $\mathcal{N}$ can be partitioned into two sets $\pi_l$ and $\pi_m$, in such a way that no two vertices in the same set are adjacent. The pair $(\pi_l, \pi_m)$ is called the *bipartition* of $\Gamma$.

Now I introduce Finite Asymmetric Contests.

**Definition 6.** A *Finite Asymmetric Contest* (FAC), is a tuple $\left\{\mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, G\right\}$, where

1. $\mathcal{N}$ is the finite set of players. $|\mathcal{N}| = N \geq 4$.

2. $\Pi = \{\pi_l, \pi_m\}$ is a partition of $\mathcal{N}$. Both $|\pi_l|, |\pi_m| \geq 2$.

3. $\mathcal{S}_\Pi = \{S_{\pi_l}, S_{\pi_m}\}$ is a collection of strategy sets with $S_i = S_{\pi_l}$ for all $i \in \pi_l$, and $S_j = S_{\pi_m}$ for all $j \in \pi_m$.

4. $\Gamma = (\mathcal{N}, E)$ is a path-connected bipartite graph with bipartition $(\pi_l, \pi_m)$.

5. $G = \left(S_{\pi_l}, S_{\pi_m}, u^{\pi_l \pi_m}, u^{\pi_m \pi_l}\right)$ is a 2-player game - the *local-interaction*.

The above primitives generate a vector of utility functions, $\mathbf{U} = \{U_1, \ldots, U_N\}$, where for any player $i \in \pi_l$, the utility from taking action $s' \in S_{\pi_l}$ is

$$U_i(s', \mathbf{s}_{-i}) := \sum_{k \in \pi_m} u^{\pi_l \pi_m}(s', s_k)$$

and for any player $j \in \pi_m$, the utility from taking action $s'' \in S_{\pi_m}$ is

$$U_j(s'', \mathbf{s}_{-j}) := \sum_{h \in \pi_l} u^{\pi_m \pi_l}(s'', s_h)$$

Again, Finite Asymmetric Contests are large population interactions, where players interact pairwise via a common *local-interaction* game. Utilities are given by the sum of payoffs earned with neighbours, where the same action must be used with each. Since the population is partitioned into two distinct groups, the local-interaction is not constrained symmetric, though of course it may be.

Definition 6 is also broad. The *Asymmetric Contests* studied in Samuelson and Zhang (1992) and Samuelson (1994) are continuum population versions of a fully connected FAC. There is also a large experimental literature that has developed related to so-called "colouring problems". In the simplest of these experiments, subjects are arranged on a bipartite graph and interact pairwise via a $2 \times 2$ game of anti-coordination. See McCubbins, Paturi, and Weller (2008) and Kearns, Suri, and Montfort (2006) for example.[19]

Having highlighted the many merits of both Group Games and Finite Asymmetric Contests, I now discuss some limitations of each.

The limitation to Group Games is exactly their strength. That the same symmetric local-interaction game occurs along every link, makes population be-

[19] If the anti-coordiantion common local-interaction is symmetric, these experiments are also Group Games, showing that the definitions are not mutually exclusive.

haviour easy to track. For example, when the local-interaction game is one of pure coordination your neighbours feel identically about local-interaction equilibrium profiles as you do, and so all players, regardless of number of neighbours, share a similar coordination problem. Furthermore, the returns from any local-interaction are dependent only on the actions chosen and not the identity of the opponent.

Finite Asymmetric Contests are useful because they allow the common local-interaction to be asymmetric. However, despite the asymmetry permitted in the game form, it is still the case that each player only interacts with those from the other group, and not with those from his/her own group. As such, similarly to Group Games, each player's payoff depends only on his choice of action and the action choices of his/her neighbours - the exact identity of his/her neighbours is unimportant.

The definition of a Multiple-Group Game, fuses both Group Games and Asymmetric Contests, retaining the merits of each, removing the limitations of each, and adding some new properties. It allows players to be distributed on any network. It allows both symmetric and asymmetric local-interaction. It allows players to interact with those from their own group (as in GGs), and with those from the other groups (as in FACs). Even in the case where all players share a common strategy set, it allows a player's local-interaction payoff functions to vary with those from different groups.

Before getting to the formal definition, some extra information is required.

For any two nonempty subsets, $A, B \subseteq \mathcal{N}$, consider the set of edges with one end in $A$ and the other in $B$. This is denoted $E^{AB}$ (which clearly equals $E^{BA}$).

Fix a graph $\Gamma = (\mathcal{N}, E)$. The undirected graph $\Gamma' = (\mathcal{N}', E')$ is said to be a *subgraph* of network $\Gamma$, if $\mathcal{N}' \subseteq \mathcal{N}$ and $E' \subseteq E$. We may consider the subgraph of $\Gamma$ with vertex set $A \subseteq \mathcal{N}$ and with edge set consisting of all edges that have both ends in $A$ (this is $E^{AA}$ as defined above). I call this the subgraph induced by $A$ and write it as $\Gamma[A]$. Similarly define a subgraph induced by a set of edges $E' \subseteq E$, and vertex set consisting of all vertices which lie at an end of an edge in $E'$. This is denoted by $\Gamma[E']$.

Lastly, note that for any set $A \subseteq \mathcal{N}$, $\Gamma[A] = \Gamma(E^{AA}) = \big(A, E^{AA}\big)$.

Let $\Pi = \{\pi_1, \ldots, \pi_M\}$ be a partition of the player set $\mathcal{N}$. By equipping the underlying graph $\Gamma = (\mathcal{N}, E)$ with this partition, $\Gamma$ can be written as

$$\Gamma = \left( \bigcup_{1 \le l \le M} \pi_l \ , \bigcup_{1 \le l \le m \le M} E^{\pi_l \pi_m} \right)$$

where $E^{\pi_l \pi_m} \cap E^{\pi_n \pi_o} = \emptyset$, if $l \neq n$ or $m \neq o$. Finally, I will say that $\Gamma$ is *partition connected according to* $\Pi$ if $\Gamma[\pi_l]$ is path connected for all $\pi_l \in \Pi$ (note this does not imply path-connectedness of $\Gamma$).

Let $\mathcal{S}_\Pi = \{S_{\pi_l}\}_{1 \le l \le M}$ be a collection of strategy sets where, for each $l$, $S_{\pi_l}$ is the strategy set common to all players in group $\pi_l$. Let $\mathbb{G}_{(\mathcal{S}_\Pi, \Gamma)} := \{G^{\pi_l \pi_m}\}_{1 \le l \le m \le M}$ be a collection of 2-player games, where $G^{\pi_l \pi_m} = \big(S_{\pi_l}, S_{\pi_m}, u^{\pi_l \pi_m}, u^{\pi_m \pi_l}\big)$ is the exchange that occurs whenever a player from $\pi_l$ interacts with a player from $\pi_m$. (If $E^{\pi_l \pi_m} = \emptyset$, then there is no such game). For $l = m$, $G^{\pi_l \pi_m}$ is constrained to be symmetric. Such a game will be referred to as a *within-group local-interaction game*. For $l < m$, $G^{\pi_l \pi_m}$ need not be symmetric. Such a game will be referred to as an *across-group local-interaction game*.

We are finally ready for the main definition of the paper, that of a *Multiple-Group Game (MGG)*.

**Definition 7.** A MGG is defined as the tuple $\mathcal{G} := \left\{\mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, \mathbb{G}_{(\mathcal{S}_\Pi, \Gamma)}\right\}$, where

1. $\mathcal{N}$ is the finite set of players. $|\mathcal{N}| = N \ge 2$.
2. $\Pi = \{\pi_1, \ldots, \pi_M\}$ is a finite partition of $\mathcal{N}$.
3. $\mathcal{S}_\Pi = \{S_{\pi_l}\}_{1 \le l \le M}$, a collection of strategy sets, one for each group $\pi_l \in \Pi$.
4. $\Gamma = (\mathcal{N}, E)$ is a path connected graph, partition connected according to $\Pi$.
5. $\mathbb{G}_{(\mathcal{S}_\Pi, \Gamma)} = \left\{G^{\pi_l \pi_m}\right\}_{1 \le l \le m \le M}$ is a collection of local-interactions.

The above primitives generate a vector of utility functions, $\mathbf{U} = \{U_i\}_{i \in \mathcal{N}}$, one for each player in the population. For player $i$ in some group, say $\pi_l$, the utility from taking action $s \in S_{\pi_l}$ is

$$U_i(s, \mathbf{s}_{-i}) := \sum_{\substack{j \in \pi_l \\ j \neq i}} g_{ij}\, u^{\pi_l \pi_l}(s, s_j) + \sum_{\substack{\pi_m \in \Pi \\ \pi_m \neq \pi_l}} \sum_{k \in \pi_m} g_{ik}\, u^{\pi_l \pi_m}(s, s_k) \tag{1.1}$$

A Multiple-Group Game is the natural extension achieved by synthesising Group Games with Finite Asymmetric Contests. It is still required that the same action must be used in all local-interactions, and that utilities are the sum of payoffs earned from each neighbour. The difference is that players may interact with those in their own group and also with those in different groups. Players may feel differently about local-interaction strategy profiles with those in the different groups (in fact, strategy profiles may not even be the same unless all groups have the same set of actions).

More concretely, if the local-interaction game in a Group Game is one of pure coordination, all players share a common preference over local-interaction profiles which, at least in the fully-connected case, can be extended to a common preference over population profiles. In other words the Group Game is itself a pure coordination problem. Whereas in a Multiple-Group Game, players may now interact with those who are in essence the same as they are (those that have a common preference over local interaction-profiles), but may now also interact with those that may have a different preference over population profiles.

It is not difficult to see that a MGG with $|\Pi| = M$ groups, in which the network is path connected and partition connected has a minimum of $2M - 1$ (1 game for every within-group interaction, and $M - 1$ across-group interactions if the groups are arranged in a "line"), and a maximum of $(\sum_{l=1}^{M} l)$ local-interaction games (1 local-interaction game for within each group, and each group interacts with $M-1$ other groups). This translates to a minimum of $3M-2$ and a maximum of $(M + 2\sum_{l=1}^{M-1} l)$ local-interaction payoff functions (1 common payoff function for each within-group game and 2 payoff functions for each across-group game).

One tractable subclass of MGGs that I think are particularly suited to experimental work are the *opponent-independent* MGGs. In these games, we have that all players share a common strategy set $S$, and for all pairs of groups $\pi_l$ and

$\pi_m$, we have that $u^{\pi_l \pi_m} = u^{\pi_l \pi_l}$. Thus, for any player $i$, from any group $\pi_l$, he/she has the same payoff function for all local-interaction profiles, regardless of who he/she interacts with. To say it more simply, player $i$ ranks *all* local-interaction profiles the same regardless of the opponent's identity.

Clearly a Multiple-Group Game is not a symmetric game unless the network is fully-connected and the partition has only one element, i.e. it can be reduced to a fully-connected Group Game. But while a fully-connected MGG may not be a symmetric game, there is clearly some aspect of symmetry to it. The utility any player receives is independent of the identities of those to whom he is neighbouring, rather it only depends on the Groups they are from and the actions taken.

One final thing: it will be convenient to order the group and the also to order the players within each group. We may thus view the strategy space $\mathbf{S}$ as $\mathbf{S}_{\pi_1}^{|\pi_1|} \times \cdots \times \mathbf{S}_{\pi_M}^{|\pi_M|}$, and view a strategy profile $\mathbf{s}$ as a tuple $(\mathbf{s}_{\pi_1}, \ldots, \mathbf{s}_{\pi_M})$, where $\mathbf{s}_{\pi_l}$ is the $|\pi_l|$-dimensional strategy vector of group $\pi_l$. A partition symmetric profile is defined as one in which all members of the same group take the same action. Clearly a symmetric profile is partition symmetric, but the opposite need not be true.[20]

**Definition 8.** Let $\Pi = \{\pi_1, \ldots, \pi_M\}$ be a partition of $\mathcal{N}$, and $\sigma$ be a permutation of $\mathcal{N}$. The permutation $\sigma$ is a *partition permutation* if the restriction of $\sigma$ to each group $\pi_m \in \Pi$, written $\sigma|_{\pi_m}$ is a permutation of $\pi_m$.

**Definition 9.** Say that the MGG, $\mathcal{G} = \left\{\mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, \mathbb{G}_{(\mathcal{S}_\Pi, \Gamma)}\right\}$, is *partition symmetric* if for all $\pi_l$ and any $i \in \pi_l$, and for all partition permutations, $\sigma$, we have that

$$U_i(s_i, \mathbf{s}_{-i}) = U_{\sigma(i)}(s_{\sigma(i)}, \mathbf{s}_{-\sigma(i)})$$

If the MGG is fully-connected, then $\mathbf{U}$ can be condensed to $\left\{U_{\pi_l}\right\}_{1 \leq l \leq M}$, since clearly a utility function is common across all players in the same group. That is, for some partition permutation $\sigma$, for all $\pi_l \in \Pi$, and all $i \in \pi_l$, $U_i = U_{\sigma(i)} = U_{\pi_l}$.

[20] In fact, clearly players from the different groups may possess totally different strategy sets and a partition symmetric profile may still exist. But clearly there must be some overlap for a symmetric profile to be possible.

**Definition 10.** Fix an MGG $\mathcal{G} := \left\{\mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, \mathbb{G}\right\}$. The Projection of $\mathcal{G}$ onto a group $\pi_l \in \Pi$, denoted $\mathcal{G}[\pi_l]$, is the group game $\mathcal{G} := \left\{\pi_l, S_{\pi_l}, \Gamma[\pi_l], G^{\pi_l\pi_l}\right\}$.

This will be useful later when examining predicted long run outcomes. It is useful to compare whether or not the predicted strategy profile of a group changes from the case where they are involved with other groups MGG, to that in which they are in isolation.

It is useful at this point to pause and take stock of just how general Definition 7 is, by observing just how many current classes of games are encompassed by the MGG framework. The following games all fit inside the Multiple-Group Game framework.

**Multiple Group Game 1** (Any 2 Player Game)**.** When $\mathcal{N} = \{1, 2\}$, $\Pi = \{\{1\}, \{2\}\}$, both $S_1$ and $S_2$ are finite sets, $g_{12} = 1$, $u^{12} = U_1$ and $u^{21} = U_2$ are both defined on $S_1 \times S_2$. This defines *any* 2-player game.

**Multiple Group Game 2** (Fully Connected Group Game, e.g. KMR)**.** When $\mathcal{N} = \{1, \ldots, N\}$, $\Pi = \{\mathcal{N}\}$, $S_i = S$ for all $i$, $g_{ij} = 1 \iff i \neq j$, $G = (2, S, u)$ is a symmetric $2 \times 2$ game. For all $i$ we have that $U_i(s_i, \mathbf{s}_{-i}) = \sum_{j \neq i} u(s_i, s_j)$.

**Multiple Group Game 3** (Network Model)**.** When $\mathcal{N} = \{1, \ldots, N\}$, $\Pi = \{\mathcal{N}\}$, $S_i = S$ for all $i$, $g_{ij} = 1 \Rightarrow i \neq j$, $G = (2, S, u)$ is a symmetric two player game. For all $i$ we have that $U_i(s_i, \mathbf{s}_{-i}) = \sum_{j \neq i} g_{ij} u(s_i, s_j)$.[21]

**Multiple Group Game 4** (Finite Asymmetric Contests)**.** When $\mathcal{N} = \{1, \ldots, N\}$, and $\Pi = \{A, B\}$, where $|A| \geq 2, |B| \geq 2$, $S_i = S_A$ for all $i \in A$, $S_j = S_B$ for all $j \in B$, $g_{ij} = 1 \iff i \in A \Rightarrow j \in B$, or $i \in B \Rightarrow j \in A$. $G = (2, S_A, S_B, u^{AB}, u^{BA})$ is a two player game. For all $i \in A$ we have that $U_i(s_i, \mathbf{s}_{-i}) = \sum_{j \in B} u^{AB}(s_i, s_j)$. Similarly For all $j \in B$ we have that $U_j(s_j, \mathbf{s}_{-j}) = \sum_{i \in A} u^{AB}(s_j, s_i)$.

[21] Note that the implication $g_{ij} = 1 \Rightarrow i \neq j$ is only in one direction, as opposed to in a fully connected group game. For example, Ellison (1993) analyses players located on a circle, such that $g_{ij} = 1$ if and only if $j \in \{i-1, i+1\} \bmod N$.

## 1.4 MGGs with Potential

A game is said to be a *potential game* (Shapley and Monderer (1996)) if the change in each player's utility from choosing a new action can be derived from a common function, referred to as the game's *potential function*.

**Definition 11.** Say that an $N$-person game $G$ is a *weighted potential game* if there exists a function $\rho : \mathbf{S} \to \mathbb{R}$ and there exist real numbers $\lambda_1, \ldots, \lambda_N$ such that, for every player $i$, and every $\mathbf{s} \in \mathbf{S}$, and all pairs of actions $s_i, s_i' \in S_i$,

$$\lambda_i \big[ U_i(s_i, \mathbf{s}_{-i}) - U_i(s_i', \mathbf{s}_{-i}) \big] = \rho(s_i, \mathbf{s}_{-i}) - \rho(s_i', \mathbf{s}_{-i}) \tag{1.2}$$

$G$ is a *potential game* if the above definition holds with $\lambda_i = 1$ for all players $i$.

If each local-interaction game of a MGG is a potential game, then the MGG inherits this property and is itself a potential game. This is true, regardless of the graph on which the players live, as stated in the following result.

**Theorem 1.** *Fix a MGG,* $\mathcal{G} := \Big\{ \mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, \mathbb{G}_{(\mathcal{S}_\Pi, \Gamma)} \Big\}$. *Suppose that for all* $1 \leq l \leq m \leq M$*, the local interaction game* $G^{\pi_l \pi_m} \in \mathbb{G}_{(\mathcal{S}_\Pi, \Gamma)}$*, is a potential game with potential function* $\rho^{\pi_l \pi_m} : S^{\pi_l} \times S^{\pi_m} \to \mathbb{R}$. *Then* $\mathcal{G}$ *is a potential game, with potential function* $\rho^\star : \mathbf{S} \to \mathbb{R}$ *defined by*

$$\rho^\star(\mathbf{s}) := \sum_{l \leq m} \sum_{1 \leq m \leq M} \sum_{(h,k) \in E^{\pi_l \pi_m}} \rho^{\pi_l \pi_m}(s_h, s_k)$$

The mountain of notation in the proof (found in Appendix 1.10.1) masks what is an incredibly simple result, whose intuition I now sketch. Consider a profile. When one player, say player $i$, independently deviates from this profile, the only change in potential occurs in local-interactions on edges that has an end at player $i$. The total change in potential is therefore just the sum of changes in potential along each of these edges.

Thus, a direct test for whether or not a MGG is a potential game can be performed by looking at the collection of local-interaction games and ignoring the underlying graph. Finally, given that all potential games have at least one equilibrium in pure strategies, the following is an immediate corollary.

**Corollary 1.** *Fix a MGG,* $\mathcal{G} := \left\{\mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, \mathbb{G}_{(\mathcal{S}_\Pi,\Gamma)}\right\}$. *Suppose that for all* $1 \leq l \leq m \leq M$, *we have that* $G^{\pi_l\pi_m} \in \mathbb{G}_{(\mathcal{S}_\Pi,\Gamma)}$ *is an potential game with potential function* $\rho^{\pi_l\pi_m} : S^{\pi_l} \times S^{\pi_m} \to \mathbb{R}$. *Then* $\mathcal{G}$ *has an equilibrium in pure strategies.*

## 1.5 Equilibrium Selection for MGGs with Potential

Evolutionary models suppose a stage game, played by a large population of players, that gets repeated indefinitely. The population is large so that reasoning correctly about what everyone else may do in future periods is nigh on impossible, and as such, it is assumed that players follow some simple updating rule, which depends on a subset of past play and generates population wide dynamics.[22]

It is assumed that the population behaviour begins at some arbitrary profile or "state", and that the population or "system" is just *let go*. When players follow deterministic updating rules, there will be path dependence, meaning that where the process begins may constrain the set of profiles at which it ultimately stops. However, if players are further assumed to occasionally experiment, then the now-noisy dynamics remove the dependence on the starting conditions, rendering it redundant. This is the key insight of Foster and Young (1990). The resulting non-empty set of equilibrium profiles that are *selected* are known as *stochastically stable equilibria*.

In this section, I apply one of the standard noisy dynamics, the Ising model dynamics of Blume (1993), to the repeated interaction. The story goes like this. Time is discrete, begins at $t = 0$, and goes forever. The current *state* of the system is the strategy profile $\mathbf{s} \in \mathbf{S} := \prod_{1\leq l\leq M} S_{\pi_l}^{|\pi_l|}$. Each period, one player is drawn randomly (uniformly) from the population and afforded the opportunity to

[22]The most widely studied dynamic is that of myopic best response as this is closest to traditional game theory.

update his/her current action. (Clearly the only feasible one-period transitions are to states that differ in a maximum of one component.) Players attempt to best respond but are constrained. Specifically, the likelihood of playing a particular action is exponentially related to its expected payoff, where a player correctly expects the behaviour of the remainder of the population to remain unchanged. To be precise, if the current population profile is $\mathbf{s}$, and player $i \in \pi_l$ is drawn, then he/she chooses action $s_i \in S_{\pi_l}$ according to the probability distribution $p_i^\beta(s_i|\mathbf{s})$, where for any $\beta > 0$,

$$p_i^\beta(s_i|\mathbf{s}) = \frac{\exp\big(\beta\, U_i(s_i, \mathbf{s}_{-i})\big)}{\sum_{s_i' \in S_{\pi_l}} \exp\big(\beta\, U_i(s_i', \mathbf{s}_{-i})\big)} \tag{1.3}$$

Despite all players in group $\pi_l$ having the same actions available to them, the relative effectiveness of one strategy over another is dependent on who a player is connected to and how he/she values that action in each local-interaction. That is, the above probability distribution is distinct for each player $i$, hence the need to index.

As $\beta \to \infty$, it is well know that player $i$'s response distribution as defined in equation (3.3) approaches that of a best response (Blume, 1993; Young, 2001), and so for any finite $\beta$, equation (3.3) is a perturbation of the best response. It is further well known that the above perturbed best response dynamic induces an irreducible and aperiodic, and hence ergodic, Markov process with transition matrix, $P^\beta$, on the *state space* $\mathbf{S}$. Due to the ergodicity of the process, $P^\beta$ has a unique stationary distribution $\mu^\beta$.

The goal is to calculate the states to which the limiting distribution, as the perturbations of the process vanish, assign positive probability. These are the stochastically stable states. That is, we wish to calculate the limiting distribution, $\mu^\star$, where

$$\mu^\star := \lim_{\beta \to \infty} \mu^\beta$$

A state $\mathbf{s}$ is said to be *stochastically stable* if $\mu^\star(\mathbf{s}) > 0$, and *uniquely stochastically stable* if $\mu^\star(\mathbf{s}) = 1$. When each local-interaction of a MGG is a potential game, under the Ising model dynamics described above, the stochastically stable states are precisely those that maximize the potential function. This is stated formally in the following Theorem whose proof is in Appendix 1.10.1.

**Theorem 2.** *Fix a MGG,* $\mathcal{G} := \left\{\mathcal{N}, \Pi, \mathcal{S}_\Pi, \Gamma, \mathbb{G}_{(\mathcal{S}_\Pi,\Gamma)}\right\}$. *Suppose that for all* $1 \leq l \leq m \leq M$, *we have that* $G^{\pi_l\pi_m}$ *is a potential game with potential function* $\rho^{\pi_l\pi_m} : S^{\pi_l} \times S^{\pi_m} \to \mathbb{R}$. *For every* $\beta > 0$, *we have that* $P^\beta$ *has the unique stationary distribution,* $\mu^\beta$, *given by*

$$\mu^\beta(\mathbf{s}) = \frac{exp\big(\beta\rho^\star(\mathbf{s})\big)}{\sum_{\mathbf{s}\in\mathbf{S}} exp\big(\beta\rho^\star(\mathbf{s})\big)} \tag{1.4}$$

*where* $\rho^\star$ *is the potential function of* $\mathcal{G}$ *as defined in Theorem 1. i.e.*

$$\rho^\star(\mathbf{s}) := \sum_{l\leq m} \sum_{1\leq m\leq M} \sum_{(hk)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_h, s_k)$$

*Furthermore, the stochastically stable states of* $\mathcal{G}$ *are those that maximize* $\rho^\star$.

An algorithm which selects the strategy profile which maximizes the potential function would therefore be a very useful object indeed. To my knowledge no such algorithm exists.

## 1.6 2 x 2 MGGs

While evolutionary models are used far and wide, typically analysis of those with normal form exchanges is restricted to cases where local-interactions are $2\times 2$ games. This makes tracking population behaviour far easier since each agent takes a binary decision. As regards MGGs, the most tractable subclass is that which I refer to as $2\times 2$ MGGs, where the population is split into two groups and all agents share a common 2-action strategy set.[23] I will denote such a game by $\mathcal{G}^{(2\times 2)}$, where the first 2 refers to the number of groups and the second 2 refers to the number of actions available to the players.

[23] This is not as restrictive as might first appear. All symmetric $2\times 2$ normal form games are potential games, and many interesting asymmetric $2\times 2$ games are too.

$2 \times 2$ games are the workhorses of game theory. They have been studied extensively and are extremely well understood.[24] They appear throughout both the theoretical and applied literatures. By limiting attention to MGGs in which all local-interactions are $2 \times 2$ games, we may borrow extensively from these literatures (see the next Section for some examples). From a *slimy modelling* perspective, I am doing nothing more than combining various classic $2 \times 2$ games into one larger strategic situation.

The population, $\mathcal{N}$, is partitioned into 2 groups, $A$ and $B$, of sizes $N^A$ and $N^B$ respectively. Ordering the groups, with $A$ first, and then ordering the players within each group, it is possible to view the population as players $1, \ldots, N^A, N^A + 1, \ldots, N^A + N^B$. All players share a common 2-action strategy set, $S := \{a, b\}$. Thus a typical $2 \times 2$ MGG, $\mathcal{G}^{(2\times2)}$, is completely characterized by the partition $\{A, B\}$ of $\mathcal{N}$, the network on which the players live, $\Gamma$, and the following three $2 \times 2$ normal form games, $G^{AA}$, $G^{AB}$, and $G^{BB}$. Recall that within-group games, $G^{AA}$ and $G^{BB}$ in this case, are constrained symmetric, whereas the across-group game, $G^{AB}$, is not.

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $\alpha_{11}, \alpha_{11}$ | $\alpha_{12}, \alpha_{21}$ |
| $b$ | $\alpha_{21}, \alpha_{12}$ | $\alpha_{22}, \alpha_{22}$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $\beta_{11}, \beta_{11}$ | $\beta_{12}, \beta_{21}$ |
| $b$ | $\beta_{21}, \beta_{12}$ | $\beta_{22}, \beta_{22}$ |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | $\gamma_{11}, \delta_{11}$ | $\gamma_{12}, \delta_{12}$ |
| $b$ | $\gamma_{21}, \delta_{21}$ | $\gamma_{22}, \delta_{22}$ |

**Figure 1.1**: $G^{AA}, G^{AB}$, and $G^{BB}$

[24] See Rapoport and Guyer (1966) and Robinson and Goforth (2005) for a full classification of $2 \times 2$ games.

For player $i \in A$ the utility from action $s$ is

$$U_i(s, \mathbf{s}_{-i}) := \sum_{\substack{j \in A \\ j \neq i}} g_{ij}\, u^{AA}(s, s_j) + \sum_{l \in B} g_{il}\, u^{AB}(s, s_l) \tag{1.5}$$

For player $k \in B$ the utility from action $s$ is

$$U_k(s, \mathbf{s}_{-i}) := \sum_{\substack{l \in B \\ l \neq k}} g_{kl}\, u^{BB}(s, s_l) + \sum_{j \in A} g_{kj}\, u^{BA}(s, s_j) \tag{1.6}$$

There are 24 symmetric cardinal $2 \times 2$ games, and 78 in total (see Rapoport and Guyer (1966); Robinson and Goforth (2005)). Since all within-group local-interactions are constrained symmetric while those across-group are not, there are therefore $24 \times 78 \times 24$ possible $2 \times 2$ MGGs. This is far too large a class to make general statements about, but while case by case studies might be the most fruitful means of study, there are still some interesting subclasses.

Symmetric $2 \times 2$ games are always potential games. So for us, $G^{AA}$ and $G^{BB}$ always admit a potential. A potential function is unique up to the addition of a constant, so *a* potential function for $G^{AA}$, $\rho^{AA}$, is given by the following:

$$\begin{array}{ll} \rho^{AA}(a,a) = \alpha_{11} - \alpha_{21} & \rho^{AA}(a,b) = 0 \\ \rho^{AA}(b,a) = 0 & \rho^{AA}(b,b) = \alpha_{22} - \alpha_{12} \end{array}$$

I now illustrate how to convert any generic $2 \times 2$ coordination game into a potential game. Recall the game $G^{AB}$. First suppose it is a coordination game where $\gamma_{aa} > \gamma_{ba}$, $\gamma_{bb} > \gamma_{ab}$ and $\delta_{aa} > \delta_{ab}$, $\delta_{bb} > \delta_{ba}$. Now define $q_A^*$, $q_B^*$, and $\zeta_A^*$, $\zeta_B^*$ as follows:

$$q_A^* := \frac{1}{(\gamma_{aa} - \gamma_{ba}) + (\gamma_{bb} - \gamma_{ab})} \tag{1.7}$$

$$q_B^* := \frac{1}{(\delta_{aa} - \delta_{ab}) + (\delta_{bb} - \delta_{ba})} \tag{1.8}$$

$$\zeta_A^* := q_A^*(\gamma_{aa} - \gamma_{ba}) \tag{1.9}$$

$$\zeta_B^* := q_B^*(\delta_{aa} - \delta_{ab}) \tag{1.10}$$

Then define a new game $\hat{G}^{AB} = (\hat{u}^{AB}, \hat{u}^{BA}, S)$, where $\hat{u}^{AB} = q_A^* u^{AB}$ and $\hat{u}^{BA} = q_B^* u^{BA}$. It can be checked that $\hat{G}^{AB}$ is a potential game with potential function $\hat{\rho}^{AB}$, where

$$\begin{array}{ll} \hat{\rho}^{AB}(a,a) = \zeta_A^* + \zeta_B^* & \hat{\rho}^{AB}(a,b) = \zeta_A^* \\ \hat{\rho}^{AB}(b,a) = \zeta_B^* & \hat{\rho}^{AB}(b,b) = 1 \end{array}$$

Thus if $\mathcal{G}^{(2\times2)}$ is a $2 \times 2$ MGG where $G^{AB}$ is a potential game, it can be transformed into a potential game $\hat{\mathcal{G}}^{2\times2}$ where each $A$ Group player's payoff is scaled by $q_A^*$ and each $B$ Group player's payoff is scaled by $q_B^*$.[25]

In the next section we discuss various $2 \times 2$ MGGs. All these games will have more than one strict equilibrium which will make the issue of equilibrium selection an interesting avenue of further research.

An *opponent-independent* MGG insists that all local interaction payoff functions are the same regardless of the opponent. In the $2 \times 2$ case, this requires that $\alpha_{ij} = \gamma_{ij}$, and $\beta_{kl} = \delta_{kl}$, for all $i, j, k, l \in \{1, 2\}$.

Of additional interest, will be those MGGs where the local-interactions $G^{AA}$ and $G^{BB}$ have *similar* but *opposite* properties.

**Definition 12.** Fix two symmetric 2×2 games $G^1 = (2, S, u^1)$ and $G^2 = (2, S, u^2)$.[26] Say that $G^1$ and $G^2$ are *mirror image* if the ordinal ranking on $S \times S$ induced by $u^1$ is opposite to that of $u^2$.[27]

[25]A similar procedure is carried out in the case that $G^{AB}$ is an *anti-coordination* game (a $2 \times 2$ anti-coordination game is a coordination game in which all players have the same 2-action strategy set and the asymmetric profiles are strict equilibria). The common deflection that an anti-coordination game can be *converted* into a game of coordination by a relabeling of strategies does not really carry weight in this case since players are already involved in other local-interactions in which the actions have a fixed interpretation.

[26]Note that the strategy sets are the same in both. This is not strictly necessary however. All we require is that both strategy set have two elements, and we identify each strategy in the first game with a corresponding one in the second game.

[27]In category theory, the opposite category or dual category, $C^{\text{op}}$, of a given category $C$ is formed by reversing the morphisms. In our simple setting, we just reverse the direction of inequalities in a partial order (a binary relation that is reflexive, anti-symmetric, and transitive). So if $(X, \geqslant)$ is a partially ordered set, we can define a new partial order on $X$, $\geqslant^*$, where for any $x, y \in X$, we have $x \geqslant^* y$ if and only if $y \geqslant x$.

**Definition 13.** Fix two symmetric $2 \times 2$ games $\Gamma^1 = (2, S, u^1)$ and $\Gamma^2 = (2, S, u^2)$. Say that $\Gamma^1$ and $\Gamma^2$ are *payoff mirror image* if they are mirror image games and the numerical values taken by $u^1$ and $u^2$ coincide.

## 1.7 Examples

This Section looks at a series of $2 \times 2$ MGGs, that describe large population situations which I claim existing models are unable to capture. Each is so simple that it could be used in a classroom experiment like that described in Section 1.2. Throughout, the population will be partitioned into two groups, Group $A$ and Group $B$, of sizes $N^A$ and $N^B$ respectively. It will make sense, and make it simple to examine the set of equilibria, to suppose that the players reside on a fully connected network. Each agent chooses one of two actions from the set $\{a, b\}$.

I will be focusing on *group-symmetric* equilibrium profiles, where all players in a given group adopt the same action. For example, $(\mathbf{a}, \mathbf{b})$ denotes that profile in which everyone in Group $A$ takes action $a$ while all those in Group $B$ adopt action $b$, etc. Payoffs when a Group $A$ player interacts with another Group $A$ player are given by $G^{AA}$, and $G^{AB}$ and $G^{BB}$ have similar interpretations. The payoffs given will be numerical - that they can be perturbed is the whole point of the paper.

### Game 1: Competing Standards

There is a population of agents each of whom chooses one of two operating systems. Due to network effects, it is desirable to have the same system as others (Katz and Shapiro, 1985; Farrell and Saloner, 1985; Arthur, 1989). However, not everybody shares the same preferences over operating systems. Group $A$ members have a leaning for coordinating on system $a$, while those in Group $B$ prefer system $b$. Furthermore, although coordination is always good, it is varyingly so - players value coordinating with those in their own group more than with those from the other group.

Payoffs that describe this situation are given in Figure 1.2 below,

The expected returns to an action are increasing in the number of other

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $4,4$ | $0,0$ |
| $b$ | $0,0$ | $2,2$ |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | $3,1$ | $0,0$ |
| $b$ | $0,0$ | $1,3$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $2,2$ | $0,0$ |
| $b$ | $0,0$ | $4,4$ |

**Figure 1.2**: Technology Adoption

agents who also adopt it. Note that the local-interactions are not opponent independent, since within-group coordination yields a higher payoff than across-group coordination. This captures the notion that coordination with those with similar preferences is more fruitful. Note further that all types of failed coordination are equally bad. It is also the case that coordinating with those in your own group on your less preferred action yields a lower payoff than coordinating with those in the other group on your more preferred action.

It is easy to show that the only candidate group-symmetric profiles for equilibria in this game are $(\mathbf{a}, \mathbf{a})$, $(\mathbf{b}, \mathbf{b})$, and $(\mathbf{a}, \mathbf{b})$. The first two of these profiles are intuitive. The profile $(\mathbf{a}, \mathbf{b})$ is interesting at it permits the possibility of more than one system emerging (even though the population is fully connected).

## Game 2: Schelling's Segregation Model

Schelling (1969, 1971) showed that people having even a small preference for the race of their neighbours could lead to complete segregation in a mixed race society. Schelling's model was graphical, with the most famous setting that in which agents are located on a circle. The game begins with agents making geographic choices of where to live from a large set of options. Once decisions have been taken, the agents are content if at least one of their two neighbours is from their own group. If both neighbours are from the other group, the agent is unhappy.

This game is repeated, so that every period one agent is afforded the option to change action. If this agent is unhappy, then the agent moves. If the agent is content, he/she stayed put.

My example is given below. Actions refer to locational decisions, though the model is not graphical in that everybody is assumed connected to everybody else.

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $4, 4$ | $0, 0$ |
| $b$ | $0, 0$ | $2, 2$ |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | $-1, -1$ | $2, 1$ |
| $b$ | $1, 2$ | $-1, -1$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $4, 4$ | $0, 0$ |
| $b$ | $0, 0$ | $2, 2$ |

Everybody agrees that venue $a$ is superior to venue $b$. In fact, the payoffs imply that each group's most preferred outcome is for them to collectively locate at venue $a$ and for the other group to collectively locate at venue $b$. For each agent from the other group who locates at the same venue as you, you experience disutility. This disutility is location independent. Since location $b$ is inferior (perhaps it is less resource rich) to location $a$, for each agent from the other group who does not locate at the same venue as you, your utility is increased by 2 when you locate at venue $a$ and by 1 when you locate at venue $b$.

There are at least 3 equilibria to the above game: $(\mathbf{a}, \mathbf{b})$, $(\mathbf{b}, \mathbf{a})$, and $(\mathbf{a}, \mathbf{a})$. The first two of these are intuitive. The third is the one I would like to discuss. In Schelling's model, all stable outcomes involve some form of segregation, and the "most reasonable" outcome involve complete segregation. But in this case it is an equilibrium for both groups to locate at venue $a$ since it is a sufficiently more attractive location. Moreover, Schelling's model is silent on how much people dislike other groups, rather they are just assumed to be content provided at least one neighbour is the same. The above setting allows the modeler to vary the intensity of preferences and to see how the equilibrium set might change as a result.

## Game 3: Schoolyard Politics

Groups $A$ and $B$ together comprise a high school class. On Friday night, each student decides to attend the arcade, $a$, or the bowling alley, $b$. Each activity is equally enjoyable but the payoffs depend on who else is also present. Each player gets a payoff of 2 for every member of their own group who attends the same venue. However, Group $A$ are "cool" and don't want to associate with the "nerds" from Group $B$. Thus, they receive a disutility for each Group $B$ member who turns up to the same venue. Conversely, Group $B$ students want to be seen with Group $A$ members, and receive a positive utility for each Group $A$ player who attends the same venue.

Payoffs are as follows,

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $2,2$ | $0,0$ |
| $b$ | $0,0$ | $2,2$ |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | $-2,2$ | $0,0$ |
| $b$ | $0,0$ | $-2,2$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $2,2$ | $0,0$ |
| $b$ | $0,0$ | $2,2$ |

This is the first situation I analyse where the across-group interaction has players with very different preferences. In fact, $G^{AB}$ is similar to the zero-sum game "matching pennies". Depending on Group Sizes, and strength of preferences, any of the four group-symmetric profiles can be equilibria to this game.

One could also imagine that the nerds in Group $B$, are eager to join with those in Group $A$ but are also eager to avoid each other. This would require altering the payoffs in $G^{BB}$ so that it is also a game of anti-coordination.

## Game 4: Community Enforcement

The prisoner's dilemma is perhaps the most famous $2 \times 2$ game. Each player has two strategies one of which is dominant. And yet, paradoxically, adopting these strategies leaves both players worse off than if they had chosen the other strategy. In a one-shot playing, it is impossible for players to cooperate on the mutually beneficial outcome, and this forms the basis for much of the repeated games literature. This literature, in particular the folk theorems,[28] show that cooperating is possible if the players are infinitely lived. Unfortunately this requires that agents are sophisticated enough to make plans long into the future.

With myopic agents who best respond, it is seemingly impossible to generate cooperation since each period is treated as a one shot playing. However, using my framework it is possible for players to take a dominated strategy in one local-interaction if the same action is not dominated in another local-interaction. Consider the following payoffs.

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $3,3$ | $0,4$ |
| $b$ | $4,0$ | $1,1$ |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | $3,3$ | $1,0$ |
| $b$ | $0,1$ | $0,0$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $3,3$ | $0,4$ |
| $b$ | $4,0$ | $1,1$ |

The story goes as follows. Players choose to make an investment in either education, $a$, or cultural identity, $b$. When interacting with those in your own group, cultural identity is dominant.[29] This is true for both groups. However, when interacting with those in the other group, investing in cultural identity is dominated. (In this regard I am envisioning the case where both groups are quite diverse.) Coordination on education is the most desired outcome, while choosing education yields a positive payoff even when matched with someone choosing cultural identity, since it allows the agent to understand those in the other group better.

[28]The number of papers in this area is vast. Some of the originals are Rubinstein (1979) and Fudenberg and Maskin (1986).

[29]There is a large literature devoted to explaining the low academic performances of certain minorities. Austen-Smith and Fryer (2005) claim that historical circumstance led to the development of a negative attitude among young blacks. As such, striving for academic achievement is deemed "acting white" and punished by exclusion from the peer group. This threat makes it dominant to invest in skills useful only within the community, rather than skills which yield returns outside.

## 1.8 Strategies, Types, and Mutations

In this Section, I claim certain terms from the biological branch of evolutionary game theory have given rise to confusion in economics.

### 1.8.1 Strategies as Types

Consider the following symmetric game commonly called Hawk-Dove. There are two actions Hawk and Dove, $\{H, D\}$. The basic premise is that while each player prefers the other to yield (i.e. choose $D$), the worst outcome occurs when both players do not yield (i.e. choose $H$).

| | $H$ | $D$ |
|---|---|---|
| $H$ | $(V-C)/2, (V-C)/2$ | $V, 0$ |
| $D$ | $0, V$ | $V/2, V/2$ |

**Figure 1.3**: The Hawk-Dove Game

The story stems from evolutionary biology and originally appeared in the classic work of Smith and Price (1973). Imagine two identical animals contesting a resource (a territory, an attractive carcass, etc.). The resource has value $V$. If both choose $D$, they can share the resource peacefully. If both choose $H$, they fight until one is seriously wounded by an amount $C$. If one chooses $D$ and the other $H$, then the one who chose $D$ retreats and leaves the prize for the one who chose $H$.

It is typically assumed that $C > V > 0$. This implies that the game has a unique symmetric equilibrium in mixed strategies where each player adopts action $H$ with probability $V/C$.[30]

Now imagine a continuum of agents randomly matched to play the above game. The only stable outcome is perfectly related to the mixed strategy equilibrium in which both players mix. More precisely, the only stable outcome is for a

[30]It is standard at this point to mention that this game also has two pure strategy equilibria in which the two players choose different strategies. But these equilibria are not relevant in an evolutionary setting as a player cannot condition behaviour on being row or column.

fraction, $V/C$, of the population to adopt action $H$, and the remaining fraction, $1 - V/C$, of the population to adopt action $D$.

A population is said to be *polymorphic* if different fractions of the agents adopt different actions. In the above setting, the situation in which $V/C$ of the population adopt action $H$, and the remainder adopt action $D$ is a stable polymorphic population.

But to me this seems somewhat of a misnomer, as it is the case that players are defined only by their actions. I am no evolutionary biologist but I am unsure if this terminology makes sense in an economic setting. For example, in economics this game is typically used to analyse common property resources. But since the game is symmetric, I would interpret a what evolutionary biologists would term a polymorphic population, as one of identical agents, some of whom are choosing different actions.

Now consider the $2 \times 2$ MGG, the Language Game (Neary, 2010a). In particular, consider the profile $(\mathbf{a}, \mathbf{a})$. This is a setting in which there are two groups of agents, with very different preferences, but they are all choosing the same action. Is this a *monomorphic* population? I would claim that it is not. For example, if the game is modeling the adoption of technological standards, then those with an underlying preference for coordinating on Mac using a Mac are (to me at least) very different to those with an underlying preference for PCs using a Mac.

Thus, I would argue that in any setting with large populations, and in particular in one with multiple large populations, the biologists' convention of classifying an agent solely by their behaviour is incorrect. I believe it would make more sense to define an agent's type as a (group, action) pair, rather that just the action they adopt.

### 1.8.2 Mutations

In an evolutionary game the modeler assigns an *object of selection*. Typically these objects of selection are actions (though, as in Dekel, Ely, and Yilankaya (2007) and Ok and Vega-Redondo (2001), they could be preferences, for example).

The object of selection is what can mutate. But this dates back to the following quote from Smith (1982) which defines an ESS (*evolutionary stable strategy*): "a strategy such that, if all members of the population adopt it, then no mutant strategy could invade the population under the influence of natural selection". Thus, if the object of selection is strategies, then mutations can only occur for these. But I believe that even if strategies are the objects of selection, by defining players by only their actions creates confusion, as it should be also possible for preferences to *change/mutate* with time. This is particularly relevant for situations with multiple populations, as the following example will hopefully demonstrate.

Again consider the Language Game. There is a finite population of agents, with each agent possessing either the preferences of Group $A$ or those of Group $B$. Local-interaction payoff functions determine the extent to which players in the different groups value coordination on the different actions. Thus, when players are myopic best responders, a mutation for a given player will mean the accidental choosing of a suboptimal action. But is it not also possible for a player's preferences to change in addition to them changing their action? Is it not possible for a player in Group $B$ to have their preferences change such that they now have preferences akin to those in Group $A$?

One could imagine a player's preferences changing such that they "mutate" across groups. I do not propose any mechanism by which this would occur, but the following is one very straightforward possibility. Suppose a player in Group $B$ has been using operating system $a$ for many periods (perhaps it is optimal given their surroundings, perhaps they have not updated in a while. etc.). Is it not possible for this player to realise the benefits of system $a$ and rearrange their working life accordingly? They might even end up with a preference for coordinating on action $a$. Thus they would become a member of Group $B$.

Thus, one scenario might be a setting in which players are myopic best responders, who occasionally mutate by choosing suboptimal actions, but also occasionally have their preferences change if they have been using the suboptimal action for many periods.

## 1.9 Conclusion

I have argued that existing large population models are constrained in that they only ever possess one local-interaction, and that this greatly limits how applicable these models can be. This paper proposes a new class of large population models, which I term Multiple-Group Games, with this one constraint removed. The population is partitioned into *groups* of like-minded agents, and it is possible that agents interact both with those in their own group **and** with those in other groups.

It is my hope that economists will find the framework useful for describing certain societal settings.

# 1.10 APPENDIX

## 1.10.1 Proofs

**Proof of Theorem 1.**

Consider the world from the perspective of some player, say $i$, in some group, say $\pi_l$. Consider two strategy profiles $\mathbf{s}, \mathbf{s}'$ which differ only in how player $i \in \pi_l$ behaves. That is, $\mathbf{s} = (s_i, \mathbf{s}_{-i})$ and $\mathbf{s}' = (s_i', \mathbf{s}_{-i})$. Suppose the current profile is $\mathbf{s}'$ and $i$ deviates and plays $s_i$. Let us examine how his/her utility changes.

$$\begin{aligned} U_i(\mathbf{s}) - U_i(\mathbf{s}') &= U_i(s_i, \mathbf{s}_{-i}) - U_i(s_i', \mathbf{s}_{-i}) \\ &= \sum_{\substack{j \in \pi_l \\ j \neq i}} g_{ij} \left[ u^{\pi_l \pi_l}(s_i, s_j) - u^{\pi_l \pi_l}(s_i', s_j) \right] \\ &+ \sum_{\substack{\pi_m \in \Pi \\ \pi_m \neq \pi_l}} \sum_{k \in \pi_m} g_{ik} \left[ u^{\pi_l \pi_m}(s_i, s_k) - u^{\pi_l \pi_m}(s_i', s_k) \right] \end{aligned}$$

Recall that if $\Gamma = (\mathcal{N}, E)$, we can decompose the set of links, $E$, in the following way

$$E = \bigcup_{1 \le l \le m \le M} E^{\pi_l \pi_m}$$

Now we exploit the fact that when player $i$ deviates, there is no change in the $j^{th}$-component of the profile for all $j \neq i$. Hence there is no local potential change on any link that does not involve player $i$. So continuing from above we have that

$$
\begin{aligned}
U_i(\mathbf{s}) - U_i(\mathbf{s}') &= \sum_{\substack{j\in\pi_l \\ j\neq i}} g_{ij}\left[\rho^{\pi_l\pi_l}(s_i,s_j) - \rho^{\pi_l\pi_l}(s_i',s_j)\right] \\
&+ \sum_{\substack{\pi_m\in\Pi \\ \pi_m\neq\pi_l}} \sum_{k\in\pi_m} g_{ik}\left[\rho^{\pi_l\pi_m}(s_i,s_k) - \rho^{\pi_l\pi_m}(s_i',s_k)\right] \\
&= \sum_{(i,j)\in E^{\pi_l\pi_l}} \left[\rho^{\pi_l\pi_l}(s_i,s_j) - \rho^{\pi_l\pi_l}(s_i',s_j')\right] \\
&+ \sum_{(i,k)\in E^{\pi_l\pi_m}} \left[\rho^{\pi_l\pi_m}(s_i,s_k) - \rho^{\pi_l\pi_m}(s_i',s_k')\right] \\
&+ \sum_{\substack{1\leq m\leq n\leq M \\ m,n\neq l}} \sum_{(h,k)\in E^{\pi_m\pi_n}} \left[\rho^{\pi_m\pi_n}(s_h,s_k) - \rho^{\pi_m\pi_n}(s_h',s_k')\right] \\
&= \sum_{(i,j)\in E^{\pi_l\pi_l}} \rho^{\pi_l\pi_l}(s_i,s_j) + \sum_{(i,k)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_i,s_k) \\
&+ \sum_{\substack{1\leq m\leq n\leq M \\ m,n\neq l}} \sum_{(h,k)\in E^{\pi_m\pi_n}} \rho^{\pi_m\pi_n}(s_h,s_k) \\
&- \sum_{(i,j)\in E^{\pi_l\pi_l}} \rho^{\pi_l\pi_l}(s_i',s_j') - \sum_{(i,k)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_i',s_k') \\
&- \sum_{\substack{1\leq m\leq n\leq M \\ m,n\neq l}} \sum_{(h,k)\in E^{\pi_m\pi_n}} \rho^{\pi_m\pi_n}(s_h',s_k') \\
&= \sum_{l\leq m} \sum_{1\leq m\leq M} \sum_{(h,k)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_h,s_k) \\
&- \sum_{l\leq m} \sum_{1\leq m\leq M} \sum_{(h,k)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_h',s_k')
\end{aligned}
$$

where the last line is equal to $\rho^\star(\mathbf{s}) - \rho^\star(\mathbf{s}')$ as desired.

**Proof of Theorem 2.**

Write $\mu$ and $P$ in place of $\mu^\beta$ and $P^\beta$. We will show that $\mu$ as defined in 3.7 is *a* stationary distribution of $P$, and then conclude that since $P$ has a *unique* stationary distribution, that it is $\mu$.

It is sufficient to show the following *detailed balance condition*.

$$\mu(\mathbf{s}')P_{\mathbf{s}'\mathbf{s}''} = \mu(\mathbf{s}'')P_{\mathbf{s}''\mathbf{s}'}, \quad \forall\, \mathbf{s}', \mathbf{s}'' \in \mathbf{S}$$

Then $\mu$ must be a stationary distribution since

$$\sum_{\mathbf{s}'\in\mathbf{S}} \mu(\mathbf{s}')P_{\mathbf{s}'\mathbf{s}''} = \sum_{\mathbf{s}'\in\mathbf{S}} \mu(\mathbf{s}'')P_{\mathbf{s}''\mathbf{s}'} = \mu(\mathbf{s}'')\sum_{\mathbf{s}'\in\mathbf{S}} P_{\mathbf{s}''\mathbf{s}'} = \mu(\mathbf{s}'')$$

It is clear that $P_{\mathbf{s}'\mathbf{s}''} = P_{\mathbf{s}''\mathbf{s}'} = 0$ unless either $\mathbf{s}' = \mathbf{s}''$, or for some $i$, $s'_i \neq s''_i$ and $s'_j = s''_j$ for all $j \neq i$.

Suppose the current state is $\mathbf{s}'$. The probability that player $i$ is selected to update his strategy is $1/N$. So,

$$\begin{aligned}
\mu(\mathbf{s}')P_{\mathbf{s}'\mathbf{s}''} = &\Big[\frac{1}{N}\frac{\exp\big(\beta\rho^\star(\mathbf{s}')\big)}{\sum_{\mathbf{s}\in\mathbf{S}}\exp\big(\beta\rho^\star(\mathbf{s})\big)}\Big] \\
&\times \left[\exp\Big(\beta\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s''_i, s'_j) + \beta \sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\, u^{\pi_l\pi_m}(s''_i, s'_k)\Big)\right] \\
&\div \left[\sum_{s_i\in S_i}\exp\Big(\beta\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i, s'_j) + \beta \sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\ u^{\pi_l\pi_m}(s_i, s'_k)\Big)\right]
\end{aligned}$$

where we ignored terms of $\rho^\star$ that do not involve links adjacent to player $i$'s vertex. To economize further on notation, define

$$\begin{aligned}
\zeta := &\frac{1}{N} \div \Bigg[\Bigg(\sum_{\mathbf{s}\in\mathbf{S}}\exp\big(\beta\rho^\star(\mathbf{s})\big)\Bigg) \\
&\Bigg(\sum_{s_i\in S_i}\exp\Big(\beta\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i, s'_j) + \beta \sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\ u^{\pi_l\pi_m}(s_i, s'_k)\Big)\Bigg)\Bigg]
\end{aligned}$$

where we note that $\zeta$ is constant for both states $\mathbf{s}'$ and $\mathbf{s}''$ since $s'_j = s''_j$ for all $j \neq i$. Thus

$$
\begin{aligned}
\mu(\mathbf{s}')P_{\mathbf{s}'\mathbf{s}''} &= \zeta\left[\exp\big(\beta\rho^{\star}(\mathbf{s}')\big)\right]\left[\exp\Big(\beta\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i'', s_j')\right.\\
&\left.+\beta\sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\, u^{\pi_l\pi_m}(s_i'', s_k')\Big)\right]\\
&= \zeta\left[\exp\beta\left(\rho^{\star}(\mathbf{s}') + \sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i'', s_j') + \sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\, u^{\pi_l\pi_m}(s_i'', s_k')\right)\right]\\
&= \zeta\left[\exp\beta\left(\sum_{l\leq m}\;\sum_{1\leq m\leq M}\;\sum_{(hk)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_h', s_k')\right.\right.\\
&\left.\left.+\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i'', s_j') + \sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\, u^{\pi_l\pi_m}(s_i'', s_k')\right)\right]\\
&= \zeta\left[\exp\beta\left(\sum_{l\leq m}\;\sum_{1\leq m\leq M}\;\sum_{(hk)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_h', s_k')\right.\right.\\
&+\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\left[u^{\pi_l\pi_l}(s_i', s_j') - \rho^{\pi_l\pi_l}(s_i', s_j') + \rho^{\pi_l\pi_l}(s_i'', s_j')\right]\\
&\left.\left.+\sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\left[u^{\pi_l\pi_m}(s_i', s_k') - \rho^{\pi_l\pi_m}(s_i', s_k') + \rho^{\pi_l\pi_m}(s_i'', s_k')\right]\right)\right]\\
&= \zeta\left[\exp\beta\left(\sum_{l\leq m}\;\sum_{1\leq m\leq M}\;\sum_{(hk)\in E^{\pi_l\pi_m}} \rho^{\pi_l\pi_m}(s_h'', s_k'')\right.\right.\\
&\left.\left.+\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i', s_j') + \sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\, u^{\pi_l\pi_m}(s_i', s_k')\right)\right]\\
&= \zeta\left[\exp\big(\beta\rho^{\star}(\mathbf{s}'')\big)\right]\left[\exp\Big(\beta\sum_{\substack{j\in\pi_l\\ j\neq i}} g_{ij}\, u^{\pi_l\pi_l}(s_i', s_j')\right.\\
&\left.+\beta\sum_{\substack{\pi_m\in\Pi\\ \pi_m\neq\pi_l}}\sum_{k\in\pi_m} g_{ik}\, u^{\pi_l\pi_m}(s_i', s_k')\Big)\right]
\end{aligned}
$$

Where the last line equals $\mu(\mathbf{s}'')P_{\mathbf{s}''\mathbf{s}'}$, and so we are done with the first part of the theorem. That the stochastically stable states are those that maximize $\rho^\star$ is immediate from the definition of $\mu$.

# Chapter 2

# Competing Conventions

**Abstract**

This paper studies a new coordination game, the Language Game, of a large but finite population. The population is partitioned into two groups of identical agents. Each player shares a common two-action strategy set and interacts pairwise with everyone else. Both symmetric profiles are pareto-efficient strict equilibria, but the groups rank them differently. The profile where successful coordination occurs only within-group, with each group adopting their most preferred action, is also an equilibrium provided the smaller group's preferences are sufficiently strong. In all dynamically stable long run outcomes, players in the same group adopt the same action. Three properties, that do not matter for equilibrium selection in the homogeneous agent models of Kandori, Mailath, and Rob (1993) and Young (1993), do matter in the Language Game. These are: group size, preference over alternative equilibria, and rates of group adaptiveness ("group dynamism"). A relative increase in group dynamism is always weakly beneficial.

"*Nobody will ever win the battle of the sexes.
There's just too much fraternizing with the enemy.*"

- Henry Kissinger

## 2.1 Introduction

A large population coordination problem is one wherein all parties can realize mutual gains, but only by making mutually consistent decisions. Often, such mutually consistent decisions require that everybody behave identically. For example: writing papers in English is a must if that is the conventional medium; it makes sense to buy a MAC if all your friends already own one; driving on the opposite side of the road as the oncoming traffic hardly seems wise; etc.

The emergence of coordinated outcomes in large societies, referred to as *conventions* by Lewis (1969) and Young (1993, 1996, 2001), has been studied using the framework of evolutionary game theory. In the canonical model, players are drawn from a homogeneous population and randomly matched to play a symmetric $2 \times 2$ game of pure coordination. This situation is then repeated with players assumed to follow some simple updating rule, that allows population behaviour to be tracked in a manageable way. However, by construction, the canonical model can only be used for studying the emergence of conventions in societies where all agents have the same preferences. This is limiting, since in many situations of interest, like the adoption of technological standards (Katz and Shapiro, 1985; Farrell and Saloner, 1985; Arthur, 1989), people often have different tastes, and so what might be best for some may not be best for all.

In this paper, I propose a new strategic situation, the "Language Game", that allows the study of conventions in a heterogeneous environment. The population is composed of not one, but two homogeneous groups. Each player has the same two-action strategy set and interacts pairwise with everybody else. I assume that successful coordination is good, while all types of failed coordination are bad. Precisely, a player's utility is linearly increasing in the number of others who adopt the same action. Each player has a most preferred coordinated outcome - the key feature is that these most preferred outcomes differ across the groups.

Uniform adoption of either action is always a strict equilibrium (convention) of the Language Game. If the smaller group has sufficiently strong preferences for one action over the other, then the profile where members of each group adopt their preferred action and forfeit coordination with those of the other group, can also be supported as an equilibrium. Such an outcome highlights one positive difference between my model and existing ones.

Next, the Language Game becomes the stage game of a repeated interaction. Time is discrete, begins at $t = 0$, and goes forever. Payoffs are received every period and actions for tomorrow must be chosen at the end of today. Following Kandori, Mailath, and Rob (1993) (hereafter KMR), players are assumed (*i*) to be *myopic*, and thus behave as though the following period is the last, and (*ii*) to exhibit some *inertia*, so that not everybody reacts instantaneously to his/her environment. Whatever individual best responses do get made in a given period are aggregated to obtain deterministic (nonergodic) group dynamics. Any decentralized adjustment process with these features will lock in on some convention with probability one.

The concept of *stochastic stability* (Foster and Young, 1990) enables crisp predictions to be made about long run behaviour. The basic idea is that players occasionally, and independently, choose non-optimal responses. These continual "mistakes" or "mutations" perturb the dynamics in such a way that population behaviour now spends the bulk of time in the neighbourhood of only certain *selected* conventions. Such equilibrium selection in the Language Game depends on three factors: group size, group payoffs, and how fast each group adapts.[1] Increased rate of adapting and increased numbers are always more likely to bring about a group's most preferred outcome. However, stronger preferences need not always be desirable.

In the canonical model with only one homogeneous population, for any convention each player's behaviour and payoff are identical. As such, the literature

[1]Rates of adapting, or "group dynamism", permits many interpretations. It can be thought of as how sensitive on average a particular group is to their surroundings, or as the frequency of death and replacement, or even as adjustment costs varying across groups since some find change less burdensome.

has focused primarily on the tension that arises when the "good" pareto-dominant equilibrium action does not coincide with the "safe" risk-dominant one. Foster and Young (1990), KMR and Young (1993) were the first to show that evolutionary forces coupled with mutations will propel population behaviour towards risk-dominance.[2] While the definition of risk-dominance, and hence of a risk dominant equilibrium, is not so clear in the Language Game, the stochastically stable equilibria are never pareto-inefficient. However, due to the existence of multiple pareto-efficient strict equilibria, not all welfare measures rank conventions identically. So inefficient outcomes may emerge although the nature of the inefficiency is different.

Ensuing work showed that the classic risk-dominance selection result is robust to situations where players interact with only small sets of neighbours, rather than with the population at large (see Ellison (1993), Blume (1993), Young (2001), and Peski (2010)). It is quite a startling finding that altering network architecture cannot in any way influence equilibrium selection.[3] However, this is simply a consequence of the fact that all pairwise interactions are the same, which implies that each individual's coordination problem is sufficiently similar, and hence in the presence of neighbours who occasionally make mistakes so is optimal behaviour, i.e. choose the risk dominant action. The Language game is a situation with more than one type of pairwise interaction, and an immediate implication of moving it to a network, is that equilibrium selection is highly sensitive to network topology. This issue is studied in a companion paper Neary (2010b).

The plan of the paper is as follows. In the next Section, I present a simple example that demonstrates how homogeneous groups with heterogeneous preferences can be a more natural way to think about certain large population situations,

[2]This negative result has generated a vast literature. The following papers show that the pareto-dominant equilibrium action can emerge in similar settings. Ely (2002), Oechssler (1999) and Oechssler (1997) are models with endogenous pairwise interactions, Canals and Vega-Redondo (1998) and Robson and Vega-Redondo (1996) vary the frequency with which players may interact, while Kim and Sobel (1995) add a round of costless communication, "cheap talk", before actions are taken.

[3]Ellison (1993) did note that network architecture can dramatically affect that speed at which selection will occur, and further notes that if selection takes a long time to occur then perhaps the validity of evolutionary forces should be called in to question.

and can provide previously unexplored insights. The Language Game is formally defined in Section 3.3, where I also characterize the set of pure strategy equilibria. Section 3.4 shows how decisions at the individual level are aggregated to yield group dynamics, and illustrates via some examples how path dependence may be influenced by the dynamics. This analysis is carried forward to Section 3.5 which contains the main results on equilibrium selection. Section 2.6 looks at welfare properties of the selected equilibria, while Section 2.7 examines how the set of selected equilibria varies as Language Game parameters change. Section 3.9 concludes and discusses some potential avenues for future work.

## 2.2 A Story

The story is an extension of one from KMR. I begin by reminding the reader of theirs and then building on it. There is a university dormitory of 10 identical students, referred to as Group $A$. Each Group $A$ student uses a computer system $s$ chosen from the set $\{a, b\}$. Each evening, the students assemble in the study hall, where everybody encounters everyone else. When two students interact, they can collaborate by playing games, sharing files, swapping add-ons, etc. But - and this is key - meetings are fruitful if and only if both students use the same computer system. Assume that system $a$ is inherently superior to system $b$. This induces the following local-interaction pure coordination game, $G^{AA}$,

| | | $A_{II}$ | |
|---|---|---|---|
| | | $a$ | $b$ |
| $A_I$ | $a$ | $2,2$ | $0,0$ |
| | $b$ | $0,0$ | $1,1$ |

where $A_I$ and $A_{II}$ are any pair of Group $A$ students. $G^{AA}$ has two pure strategy equilibria, $(a, a)$ and $(b, b)$, and a third equilibrium in mixed strategies, where each player puts weight 1/3 on $a$. The population coordination problem has two pure strategy equilibria in which all 10 students adopt a common system, $a$ or $b$. These profiles are denoted by $\mathbf{a}$ and $\mathbf{b}$ respectively.

It is assumed, again following KMR, that students occasionally have the opportunity to change their computer, and that students are myopic in that decisions are taken based on the current distribution of computers. This generates *Darwinian* dynamics, in which population behaviour is always *drifting* towards either **a** or **b**. Initial conditions are key: if more than one third of the population begins using system $a$ (4 or more since the population is of size 10), then outcome **a** will be reached, while if 7 or more students start out using $b$, then **b** will be the final resting point. The reasoning is simple, all players collectively agree on what action is a best response, so the best response today must be at least as good a response tomorrow as the number of players taking that action can only (weakly) increase.

However, when *trembles* or *mistakes* or *experimentation* are incorporated into the dynamics, it is possible to *select* between strict equilibrium outcomes. Suppose that the probability that a student mistakenly chooses the computer that is not an optimal response is given by $\varepsilon$. It takes 4 or more simultaneously occurring mistakes to dislodge the system from **b**, and 7 or more to get away from **a**. The most likely events of this form occur with probability of orders $\varepsilon^4$ and $\varepsilon^7$ respectively. For small values of $\varepsilon$, $\varepsilon^7 \ll \varepsilon^4$, and so KMR conclude that when agents are myopic best responders, who occasionally make mistakes, that outcome **a** is far more likely to be observed in the long run.

A key component of the above story was that system $a$ is *inherently superior* to system $b$. While in many coordination problems it is plausible to believe that coordination on one particular strategy is better (by any metric) than another, words like "better" derive from primitive preferences, and preferences are individual by nature. In a population with heterogeneous agents, what might be best for some may not be best for all.

To illustrate the impact of adding heterogeneity, consider the following extension to the above story. Suppose Group $A$ are "slackers" - they must also do coursework on their machines, but their main use for computers is playing games. Instead of assuming that computer system $a$ is flat out better than $b$, let us suppose that system $a$ more readily supports gaming platforms, which justifies Group

$A$'s underlying preference for coordinating on it. Suppose further that there is another dorm of 5 more students in the next building. This dorm is unconnected to the first dorm, and I shall call those in this dorm, Group $B$. Every night these 5 Group $B$ students meet in a separate study hall and exchange software, etc. Again, interactions are beneficial if and only if the pair involved share the same system. However, those in Group $B$ are more "serious" scholars, and platform $b$ suits their scholarly needs better. The local-interaction between two Group $B$ students is given by the following pure coordination game, $G^{BB}$,

| | | $B_{II}$: $a$ | $b$ |
|---|---|---|---|
| $B_I$ | $a$ | $1,1$ | $0,0$ |
| | $b$ | $0,0$ | $2,2$ |

where $B_I$ and $B_{II}$ are any two Group $B$ students. By an identical analysis to that given above for the Group $A$ coordination problem, left to their own devices Group $B$ will adopt computer system $b$ in the long run.[4] Now, consider the 15 person population as a whole. Writing a group-symmetric profile[5] as a vector, $(\mathbf{s}_A, \mathbf{s}_B)$, with the Group $A$ profile written first, there are 4 strict equilibria: $(\mathbf{a}, \mathbf{a}), (\mathbf{a}, \mathbf{b}), (\mathbf{b}, \mathbf{a}), (\mathbf{b}, \mathbf{b})$. One can easily see, and it is quite intuitive, that the long run population profile will be $(\mathbf{a}, \mathbf{b})$. So with both groups isolated, each group internally coordinates on its preferred outcome.

The situation changes when the dorms are connected. Suppose that in an effort to free up space, the university stipulates that both groups should study in the larger Group $A$ study room which can accommodate 5 extra bodies. (This frees up the smaller study room for other activities.) So now, all 15 students meet in the same room every evening. Everybody interacts with everybody else, and within-group local-interactions are as before. It remains to specify the local-interaction that occurs when students from opposite groups meet. This is described by the

[4]The probability of Group $B$ transitioning from $\mathbf{a}$ to $\mathbf{b}$ occurs with probability of order $\varepsilon^2$, while that of transitioning from $\mathbf{b}$ to $\mathbf{a}$ occurs with probability of order $\varepsilon^4$.

[5]A *group-symmetric* profile is one in which those in the same group take the same action. A symmetric profile is group-symmetric, though plainly the reverse need not be true.

coordination game $G^{AB}$, in which the row player, $A_i$, is from Group $A$, while the column player, $B_j$, is a Group $B$ student,

| | | $B_j$ $a$ | $B_j$ $b$ |
|---|---|---|---|
| $A_i$ | $a$ | $2,1$ | $0,0$ |
| | $b$ | $0,0$ | $1,2$ |

$G^{AB}$ is not symmetric. It has two pareto-efficient pure strategy equilibria, $(a, a)$ and $(b, b)$, over which players' preferences disagree. Considering the new coordination problem, interactions are now occurring both within- and across-group. Each Group $A$ student interacts with 9 fellow Group $A$ students and 5 Group $B$ students, while each Group $B$ student interacts with 4 other Group $B$ students and 10 Group $A$ students. The only group-symmetric equilibria to this new situation are $(\mathbf{a}, \mathbf{a})$ and $(\mathbf{b}, \mathbf{b})$.[6] While both are pareto-efficient, the 10 Group $A$ students prefer $(\mathbf{a}, \mathbf{a})$, while the 5 Group $B$ students view $(\mathbf{b}, \mathbf{b})$ as most desirable.

I now pause and ask the reader to predict what they think the long run outcome will be (recalling that $(\mathbf{a}, \mathbf{a})$ and $(\mathbf{b}, \mathbf{b})$ are the only viable candidates). One conjecture might be the following. Even though behaviour evolves in a decentralised manner via individual best responses, group preferences are *mirrored* so the greater Group $A$ numbers should somehow collectively force its preferred outcome, $(\mathbf{a}, \mathbf{a})$, onto the population at large.

It turns out that the answer is subtle and depends on a variety of factors aside from group sizes and group preferences. The first of these is mistakes. In the original story with only Group $A$, all agents were identical so assuming they all make time- and profile-independent mistakes with equal probability seemed not completely unreasonable. However, with a population composed of two types of agents, if Group $A$ students tremble with probability $\varepsilon^A$ while Group $B$ students tremble with probability $\varepsilon^B$, there is no obvious reason to conclude that $\varepsilon^A = \varepsilon^B$.

[6]While $(\mathbf{a}, \mathbf{b})$, and $(\mathbf{b}, \mathbf{a})$ were equilibria when the two groups were unconnected, they are no longer equilibria when all students study together.

The second complicating factor is a property I refer to as *group dynamism*. Group dynamism can be thought of as the rate at which a group responds. It may be that Group $A$ students are more lethargic than Group $B$ students. Perhaps on average only one Group $A$ student updates his action in a given period, whereas all Group $B$ students update their action every period. In the real world, there is no reason to suppose that different groups respond at identical rates. In fact, often they do not.

However, with these caveats in mind, let's begin by assuming that: (1) payoffs are as given in $G^{AA}$, $G^{AB}$ and $G^{BB}$; (2) the group sizes are 10 and 5 for Groups $A$ and $B$ respectively; (3) mistakes are such that $\varepsilon^A = \varepsilon^B = \varepsilon$; and (4) each period, both groups evolve according to the best-reply dynamic in which all students react. Now, let us calculate how easy it is to dislodge the population from each symmetric profile.

Suppose first that the population profile is $(\mathbf{b}, \mathbf{b})$. Any Group $B$ student needs to see a minimum of 10 others taking action $a$ in order to switch his/her action, while a Group $A$ student needs to see a minimum of 6. Let's say between 6 and 9 students accidentally chose action $a$ (it is not important how these students are distributed across the two groups). In the following period, all Group $B$ students choose action $b$, but all Group $A$ students choose action $a$, so that the new profile is $(\mathbf{a}, \mathbf{b})$. With no further mistakes, all students take action $a$ the following period. The conclusion is that 6 or more simultaneous mistakes are sufficient to shift the population from $(\mathbf{b}, \mathbf{b})$ to $(\mathbf{a}, \mathbf{a})$.

Now assume that the current profile is $(\mathbf{a}, \mathbf{a})$. Payoffs are mirrored so the 6-player and 10-student bounds are still relevant. If between 6 and 9 mistakes occur whereby students accidentally choose action $b$, it is enough to induce the 5 Group $B$ students to take action $b$, but not enough to induce the 10 Group $A$ students to do so. Thus, next period's profile is $(\mathbf{a}, \mathbf{b})$, and with no further trembles, the system reverts to $(\mathbf{a}, \mathbf{a})$. It can be computed that a minimum of 10 simultaneous mistakes are required to shift the system from $(\mathbf{b}, \mathbf{b})$ to $(\mathbf{a}, \mathbf{a})$.

Transitioning from $(\mathbf{a}, \mathbf{a})$ to $(\mathbf{b}, \mathbf{b})$ requires an event that occurs with probability of order $\varepsilon^{10}$, while transitioning from $(\mathbf{b}, \mathbf{b})$ to $(\mathbf{a}, \mathbf{a})$ one with probability of order $\varepsilon^6$. For small values of $\varepsilon$, the second transition is far more likely. So provided that: (*i*) everybody interacts with everybody else; (*ii*) both groups respond according to best-reply dynamics; (*iii*) the probability of a student making a mistake is small, equal for all students, and independent of the current population profile; and (*iv*) payoffs and group-sizes are as specified above. Then, the informal analysis concludes that the unique long run outcome will be $(\mathbf{a}, \mathbf{a})$.

Now suppose that the groups adapt at different rates. Precisely: (1) payoffs are as given in $G^{AA}$, $G^{AB}$ and $G^{BB}$; (2) the group sizes are 10 and 5 for Groups $A$ and $B$ respectively; (3) mistakes are such that $\varepsilon^A = \varepsilon^B = \varepsilon$; and (4) each period, Group $B$ evolves according to the best-reply dynamic, while only one Group $A$ student best responds. Again let us calculate how easy it is to dislodge the population from each of the equilibria. The bounds of 6 and 10 are derived from preferences and group sizes, not dynamics, so those have not changed. The difference in this analysis is that it will matter exactly who is making mistakes.

This time, start with the population at $(\mathbf{b}, \mathbf{b})$. Suppose 6 Group $A$ students mistakenly choose $a$ (it does matter that these 6 students are from Group $A$). Next period, Group $B$ students maintain taking action $b$, while one more Group $A$ student adopts action $a$, so that the total number taking action $a$ is increased to 7. The following period 8 Group $A$ students are using action $a$, and so on. Once all 10 Group $A$ students are taking action $a$, action $a$ becomes optimal for Group $B$ students who all immediately adopt it. Thus, 6 of the "right kind" of mistakes are enough to transition from $(\mathbf{b}, \mathbf{b})$ to $(\mathbf{a}, \mathbf{a})$.

Now let the current profile be $(\mathbf{a}, \mathbf{a})$, and suppose that 6 Group $A$ students accidentally choose action $b$, (again, it matters that these 6 students are from Group $A$). At this new profile, Group preferences disagree. The reactiveness of the groups means that the following period, all 5 Group $B$ students adopt action $b$, while one of the 6 Group $A$ students who mistakenly chose action $b$ reverts back to action $a$. Thus, there are 10 $\big(= 5 + (6 - 1)\big)$ students taking action $b$. This is enough for all Group $B$ students to maintain action $b$ and for Group $A$ students to prefer action $b$. From this new situation, by an analysis similar to the previous paragraph, the behaviour of the population moves incrementally to $(\mathbf{b}, \mathbf{b})$.

Under these different dynamics, transitioning from $(\mathbf{b}, \mathbf{b})$ to $(\mathbf{a}, \mathbf{a})$ still occurs with probability of order $\varepsilon^6$. More importantly however, the likelihood of transitioning from $(\mathbf{a}, \mathbf{a})$ to $(\mathbf{b}, \mathbf{b})$ has been lowered to $\varepsilon^6$. So provided that: (*i*) everybody interacts with everybody else; (*ii*) Group $B$ best-replies, while Group $A$ is more lethargic; (*iii*) the probability of a student making a mistake is small, equal for all students, and independent of the current state; and (*iv*) payoffs and group-sizes are as specified above. Then, the informal analysis concludes that both symmetric outcomes are equally likely to be observed in the long run.

So what, the reader might ask, is the point of this section? There is certainly no hint of a crisp result like the risk-dominance prediction of KMR and Young (1993).[7] But in fact, the lack of a clean result is precisely the point. That is, what prediction results I do obtain, are incredibly fragile. That group size and strength of payoffs affect long run behaviour is intuitive but not predicted by homogeneous agent case. In that framework, once an equilibrium is risk-dominant it is selected (even though it may be pareto-dominated to an arbitrary extent), and this is of course independent of population size.

Once we move to the details of the dynamics the situation becomes even worse. It is well known (Bergin and Lipman, 1996) that any strict equilibrium may be selected with appropriately defined mutations. However, the risk-dominance result is robust to both uniform errors and also payoff dependent errors (Blume, 1993). This is not the case for the Language Game described above - though given that strength of payoffs affect selection even under uniform errors, it is unsurprising that payoff dependent dynamics yield different selection results.

[7] I am applying existing selection techniques to a new strategic situation, and in this situation it is unclear how concepts like risk-dominance ought be defined. One might be tempted to suggest that the profile $(\mathbf{a}, \mathbf{b})$ is risk-dominant, but for the parameters given above this profile is not an equilibrium. Thus we would be in the unusual situation of having the risk-dominant profile be unstable.

## 2.3 The Language Game

### 2.3.1 The Model

The Language Game, $\mathcal{G}$, is defined as the tuple $\{\mathcal{N}, \Pi, S, \mathbb{G}\}$, where $\mathcal{N} := \{1, \ldots, N\}$ is the *population* of players; $\Pi := \{A, B\}$ is a *partition* of $\mathcal{N}$ into two nonempty homogeneous *groups* $A$, $B$ of sizes $N^A$, $N^B$ $(\geq 2)$ respectively; $S := \{a, b\}$ is the set of actions common to all players; $\mathbb{G} := \left\{G^{AA}, G^{AB}, G^{BB}\right\}$ is the collection of pairwise local-interactions, where $G^{AA}$ is the exchange that occurs whenever a player from Group $A$ meets a player from Group $A$, etc. $G^{AA}, G^{AB}$, and $G^{BB}$ are given as follows,

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $p, p$ | $0, 0$ |
| $b$ | $0, 0$ | $1-p, 1-p$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $1-q, 1-q$ | $0, 0$ |
| $b$ | $0, 0$ | $q, q$ |

$G^{AB}$

| $A_i$ \ $B_j$ | $a$ | $b$ |
|---|---|---|
| $a$ | $p, 1-q$ | $0, 0$ |
| $b$ | $0, 0$ | $1-p, q$ |

The Language Game is a simultaneous move game, in which players do not randomize. Utilities are the sum of payoffs earned from playing the field, where the same action must be used with one and all.[8] I assume that $p, q \in (1/2, 1)$, so Group $A$ members prefer to coordinate on $a$, and Group $B$ prefer to coordinate on $b$. Even though within-group local-interactions are constrained symmetric while those across-group are not, note that all local-interactions are *opponent independent* in that a player's payoff depends only on the actions chosen and not the other player's

[8]While this has a different interpretation to a game of random matching, strategic behaviour is the same in both.

identity.[9] Thus a player cares only about the number of others expected to choose the same action, and not on who those others are.

With only two types of agents, population behaviour can be written concisely. Let $[\omega]_K$ denote the number of players in group $K \in \Pi$ using action $a$. Call $\omega = \big([\omega]_A, [\omega]_B\big)$ the *state* of the play, where the *state space* is $\Omega := \{0, \ldots, N^A\} \times \{0, \ldots, N^B\}$. For any state $\omega \in \Omega$, define $n_a = [\omega]_A + [\omega]_B$, and $n_b = N - n_a$. The utility a player in group $K \in \Pi$ receives from taking action $s \in \{a, b\}$ in state $\omega$, written $U^K(s; \omega)$, is given by

$$U^A(a;\omega) := \Big(n_a - 1\Big)p \tag{2.1}$$

$$U^A(b;\omega) := \Big(N - n_a - 1\Big)(1-p) \tag{2.2}$$

$$U^B(a;\omega) := \Big(n_a - 1\Big)(1-q) \tag{2.3}$$

$$U^B(b;\omega) := \Big(N - n_a - 1\Big)q \tag{2.4}$$

Before discussing individual behaviour, I should mention genericity. Letting $\mathbb{N} := \{1, 2, \ldots\}$, the set of Language Games can be parameterized by $\Theta = \{(N^A, N^B, p, q) : N^A, N^B \in \mathbb{N}\backslash\{1\}; p, q \in (1/2, 1)\}$. For a given game $\mathcal{G} \in \Theta$ and a given group $K \in \Pi$, the statement "if there does not exist a state $\omega \in \Omega$ such that $U^K(a;\omega) = U^K(b;\omega)$", will be abbreviated to "*genK*". If there exists such a state, the shorthand is "*ngenK*". The subset of the parameter space for which any indifference occurs can easily be shown to have a closure of measure zero, and so when both genA and genB, following standard terminology I say $\mathcal{G}$ is *generic*. Otherwise, $\mathcal{G}$ is *nongeneric*.

### 2.3.2 Individual Behaviour

Let $\mathbb{R}$ denote the real line, and $\mathbb{R}_+$ its positive part. For any $x \in \mathbb{R}$, let $\lceil x \rceil := \min\{n \in \mathbb{N} \,|\, x \leq n\}$ and $\lfloor x \rfloor := \max\{n \in \mathbb{N} \,|\, x \geq n\}$. Define the follow-

[9]This is obviously a gross simplification and may not be applicable for many real world situations. See Neary (2010c) for extensions of this model to situations where not only the value of successful coordination may be opponent dependent, but even what might be the optimal action with one Group may not be optimal with another, e.g. coordination 'vs' anti-coordination.

ing,[10]

$$n_a^A := \min\left\{n_a \,\middle|\, U^A(a;\omega) > U^A(b;\omega)\right\} = \lceil (1-p)N + (2p-1)\rceil, \tag{2.5}$$

$$n_a^B := \min\left\{n_a \,\middle|\, U^B(a;\omega) > U^B(b;\omega)\right\} = \lceil q(N-2)+1\rceil, \tag{2.6}$$

$$n_b^A := \min\left\{n_b \,\middle|\, U^A(a;\omega) < U^A(b;\omega)\right\} = \lceil p(N-2)+1\rceil, \tag{2.7}$$

$$n_b^B := \min\left\{n_b \,\middle|\, U^B(a;\omega) < U^B(b;\omega)\right\} = \lceil (1-q)N + (2q-1)\rceil \tag{2.8}$$

$n_a^A$ is the number of players taking action $a$ for a player from Group $A$ to strictly prefer action $a$, etc. When $p = q$, by symmetry $n_a^A = n_b^B$ and $n_b^A = n_a^B$. This appears to suggest that the strategic situation is mirrored for groups $A$ and $B$. While this is true at the individual level, it need not be true at the level of the group.

Let $\Omega^{A,a\succ b}$ and $\Omega^{B,a\succ b}$ denote the set of states such that $A$ players and $B$ players respectively *strictly* prefer action $a$ to action $b$. Similarly define $\Omega^{A,b\succ a}$, and $\Omega^{B,b\succ a}$.

$$\Omega^{A,a\succ b} := \left\{\omega \in \Omega \,\middle|\, [\omega]_A + [\omega]_B \geq n_a^A\right\} \tag{2.9}$$

$$\Omega^{B,a\succ b} := \left\{\omega \in \Omega \,\middle|\, [\omega]_A + [\omega]_B \geq n_a^B\right\} \tag{2.10}$$

$$\Omega^{A,b\succ a} := \left\{\omega \in \Omega \,\middle|\, [\omega]_A + [\omega]_B \leq N - n_b^A\right\} \tag{2.11}$$

$$\Omega^{B,b\succ a} := \left\{\omega \in \Omega \,\middle|\, [\omega]_A + [\omega]_B \leq N - n_b^B\right\} \tag{2.12}$$

Sets $\Omega^{A,a\succeq b}$, $\Omega^{B,a\succeq b}$, $\Omega^{A,b\succeq a}$, and $\Omega^{B,b\succeq a}$ are defined likewise but for *weak* preference. Generically, these sets of weak- and strict-preference coincide. Letting $\subseteq$ ($\subset$) denote weak (strict) inclusion, we have the following lemma whose simple proof is omitted.

**Lemma 1.** *When $N^A \geq 2$ and $N^B \geq 2$,*

1. $\Omega^{B,a\succeq b} \subseteq \Omega^{A,a\succ b}$
2. $\Omega^{A,b\succeq a} \subseteq \Omega^{B,b\succ a}$

[10] The values given are for generic $\mathcal{G}$. For the relevant nongeneric case, increase the values by 1.

An immediate implication of Lemma 1 is that, even non-generically, there is no state such that members of both groups are simultaneously indifferent, i.e. ngenA and ngenB cannot both hold.

The interpretation of the lemma is as follows. Fix a state $\omega$. If all members of Group $B$ ($A$) weakly prefer action $a$ ($b$) at this particular state, then this same action is the unique best response for all members of Group $A$ ($B$), and hence is a best response for the population as a whole. It does not say that if action $a$ ($b$) is preferred by Group $A$ ($B$), it must simultaneously be preferred by Group $B$ ($A$). That is, there may exist a state such that group preferences differ. The following provides mild sufficient conditions for the existence of such a state.

**Lemma 2.** *If either*

- *$N$ is even, or*
- *$N$ is odd and $N > 2 + \frac{1}{p+q-1}$.*

*then,*

$$\Omega^{A,a \succ b} \cap \Omega^{B,b \succ a} \neq \emptyset$$

*Proof.* The proof is found in Appendix 2.9.3. □

### 2.3.3 Equilibria

Behaviour at states $(0,0),(0,N^B),(N^A,0)$, and $(N^A,N^B)$, is referred to as *group-symmetric* for obvious reasons. These four states will appear repeatedly throughout the paper, and are denoted by $\omega_{bb}, \omega_{ba}$, $\omega_{ab}$, and $\omega_{aa}$ respectively. For a given game, $\mathcal{G}$, let $E(\mathcal{G})$ denote the set of group-symmetric equilibria. With an abuse of terminology, I will refer to $E(\mathcal{G})$ as the *equilibrium set*, since it turns out (Section 3.4) that group-symmetric equilibria are the only serious candidates for long run behaviour. The following Theorem, stated without proof, classifies $E(\mathcal{G})$ for various parameters.

**Theorem 3.** *In the Language Game, $\mathcal{G}$,*

1. *State $\omega_{aa}$ is always a strict equilibrium.*

*2. State $\omega_{bb}$ is always a strict equilibrium.*

*3. State $\omega_{ba}$ is never an equilibrium.*

*4. State $\omega_{ab}$ is an equilibrium if and only if*

$$p \geq \frac{N^B}{N-1} \quad \textit{and} \quad q \geq \frac{N^A}{N-1}$$

Parts 1 - 2 of Theorem 3 are easily understood, since deviating from a symmetric profile means failing to coordinate with everyone in the population. Part 3 is also very simple. One of the groups must be (weakly) smaller, and at state $\omega_{ba}$, members of this (weakly) smaller group observe strictly more than half the players in the population adopting their preferred action. Hence they wish to deviate.

The intuition for part 4 is as follows. State $\omega_{ab}$ involves each player successfully coordinating on their most preferred action with only those from his/her own group. To sustain $\omega_{ab}$ as an equilibrium, the high payoffs earned from within-group interactions, must exceed those that could be earned from successful coordination with the members of the other group on a less preferred action. This requires the product of "own group size" and "preferred local-interaction payoff" be sufficiently large for each player. That is, a player must either be part of the larger group, or part of a group with a strong relative preference for one action over the other, or both. The inequality for the larger group clearly always holds, and so one must only check that of the smaller group.

While Theorem 3 is simple, it is also intuitive. The following two examples, which are carried throughout the paper, illustrate precisely why. They further demonstrate that while both symmetric profiles are socially efficient, members of the different groups prefer different ones. This observation stimulates the discussion of welfare in Section 2.6.

**Example 1.** Let $\mathcal{G}^1 = (10, 5, 3/5, 2/3)$.
For these parameters $\omega_{ab}$ is not an equilibrium, because from this state Group $B$ players have an incentive to deviate to action $a$. Precisely, the second inequality of

part 4 in Theorem 3 does not hold. The reader can check that the first inequality of part 4 in Theorem 3, relevant to Group $A$ players, does hold, as it must since Group $A$ is the larger group.

**Example 2.** Let $\mathcal{G}^2 = (10, 5, 3/5, 5/6)$.
Group $B$ members now have a stronger relative preference for coordinating on action $b$ over action $a$. As in Example 1, the first inequality of Theorem 3 part 4, for $\omega_{ab}$ to be an equilibrium holds (there is no need to recheck as parameters relevant to Group $A$ are the same as they were in $\mathcal{G}^1$). This time however, the inequality relevant to Group $B$ members also holds. While group $B$ still has only half as many members as group $A$, the relative reward for coordinating on action $b$ over action $a$ for group $B$ members has increased sufficiently that even coordinating with a small number of players on their most preferred action can compensate for the larger number of *failed* coordinations.

## 2.4 Evolutionary Dynamics

### 2.4.1 Specification

Now suppose the Language Game becomes the stage game of a repeated interaction. Time is discrete, begins at $t = 0$, and goes forever. Utilities are received every period and actions for tomorrow are chosen at the end of today. I avail of precisely the assumptions placed on individual behaviour from KMR's evolutionary model.

**Assumption 1.** *Inertia: At the end of each period, a nonempty subset of players are provided with the opportunity to revise their strategy for the following period.*

**Assumption 2.** *Myopia: When a player does react, he best responds to the current environment.*

One possible explanation put forward for high inertia is that in many situations changing an action is a costly excercise. Another is that players observe only slices of information so their understanding of the game may be imperfect, and that this may cause them to stand by the status quo for longer than might be optimal. Myopia usefully captures the notion that players are boundedly rational and/or do not understand the dynamics of the population at large. Furthermore, it follows quite naturally for systems with high inertia, since in this case, a best response against the current population profile should not only generate a high payoff tomorrow, but also a "pretty good" payoff for some time in near the future.

Aggregating responses from the individual level to the population level, the evolution of population behaviour may be described by a deterministic dynamic, $\Psi$, where

$$\begin{aligned}\omega_{t+1} &:= \Psi(\omega_t)\\ &= \big(\Psi^A(\omega_t), \Psi^B(\omega_t)\big)\end{aligned}$$

The mappings $\Psi^A : \Omega \to \{0, \ldots, N^A\}$ and $\Psi^B : \Omega \to \{0, \ldots, N^B\}$ are the respective Group dynamics. Following from myopia, strategic decisions are made by looking back at today's environment, and making a choice of action for tomorrow, based on what would have been an ideal strategy to have held earlier today. However, since there is inertia, perhaps not all agents myopically best respond every period. Thus, each of $\Psi^A$ and $\Psi^B$ possess the "*Darwinian*" property of KMR, so that $\Psi$ satisfies the following definition.

**Definition 14** (Group-Darwinian Adjustment Process)**.** Say that $\Psi = (\Psi^A, \Psi^B)$ has the *Group-Darwinian Property* if, for any $K \in \{A, B\}$,

1. for all $\omega \notin \{\omega' \mid [\omega']_K = 0, N^K\}$,

$$\operatorname{sign}\big(\Psi^K(\omega) - [\omega]_K\big) = \operatorname{sign}\big(U^K(a;\omega) - U^K(b;\omega)\big) \tag{2.13}$$

2. 
   - $\Psi^K(\omega) = 0$, if $[\omega]_K = 0$ and $U^K(a;\omega) \leq U^K(b;\omega)$
   - $\Psi^K(\omega) = N^K$, if $[\omega]_K = N^K$ and $U^K(a;\omega) \geq U^K(b;\omega)$

Group-Darwinianism naturally extends the Darwinian property of KMR to a situation with multiple groups. It is *similar*, in that it makes the standard evolutionary assumption that better strategies are no worse represented next period.

But it is *different* since, $(i)$ for certain states, groups may disagree over which strategies are "better" (Lemma 2), and $(ii)$ the rate at which each group's best responses are better represented next period is left unspecified, and need not be the same across groups.

As time proceeds the population reacts every period, and so interest lies in repeated applications of $\Psi$. For all $\omega$, let $\Psi^0(\omega) = \omega$, and for $m \in \mathbb{N}$ define the $m$-fold repetition of $\Psi$, $\Psi^m$, inductively as $\Psi^m(\omega) = \Psi\big(\Psi^{m-1}(\omega)\big)$. Define the set of *rest points*, $\Omega_0 := \{\omega \,|\, \Psi(\omega) = \omega\} \subset \Omega$. It can easily be shown that for each $\omega \in \Omega$, there exists $\hat{\hat{m}}(\omega)$, such that for all $m \geq \hat{\hat{m}}(\omega)$, $\Psi^m(\omega) \in \Omega_0$. Finally, let $\hat{m} := \max_\omega \hat{\hat{m}}(\omega)$, and for each $\omega \in \Omega_0$, define the *basin of attraction* of $\omega$ by $\mathcal{V}(\omega) := \{\omega' \,|\, \forall\, m \geq \hat{m}, \Psi^m(\omega') = \omega\}$. Generically, $\Omega_0 = E(\mathcal{G})$, and so the state space can be partitioned into $\{\mathcal{V}(\omega)\}_{\omega \in E(\mathcal{G})}$. When ngenA, it is possible that $(n_a^A - 1, 0) \in \Omega_0$, and when ngenB, it may be that $(N^A, N^B - n_b^B + 1) \in \Omega_0$.

When Lemma 2 holds, the basins of attraction depend on the exact specification of $\Psi$. Keeping track of these is of prime concern for issues of equilibrium selection in Section 3.5. To assist in this, define a partial ordering on $\Omega$ as follows. If $\omega$ and $\omega'$ are elements of $\Omega$, write $\omega \geqslant_a \omega'$, if $[\omega]_A \geq [\omega']_A$ and $[\omega]_B \geq [\omega']_B$. That is, $\omega \geqslant_a \omega'$ if, in state $\omega$, there are (weakly) more players in both groups taking action $a$. The pair $(\Omega, \geqslant_a)$ is a complete lattice with *bottom* element $\omega_{bb}$ and *top* element $\omega_{aa}$.[11]

While the Group-Darwinian property of Definition 14 seems appealing at first, the class of dynamics satisfying it is still too broad to show general results. The following additional constraint placed on the dynamics will make tracking of population behaviour easier.

**Definition 15** (Monotonic Adjustment Process)**.** Say that $\Psi : \Omega \to \Omega$ is *monotonic*, if for any pair $\omega', \omega'' \in \Omega$,

$$\omega' \geqslant_a \omega'' \Rightarrow \Psi(\omega') \geqslant_a \Psi(\omega'') \tag{2.14}$$

It is easy to construct adaptive processes satisfying Definition 14, but not

[11] Lattices are briefly discussed in Appendix 2.9.1.

Definition 15, and vice versa. I limit attention to the class of dynamics satisfying both Definitions, and refer to these to as *monotone Group-Darwinian* processes. Soon I place further restrictions on this class, but first, a useful result.

**Lemma 3.** *For any monotonic Group-Darwinian process, and any state $\omega \in \Omega_0$, the set $\mathcal{V}(\omega)$ is convex.*

*Proof.* The proof is contained in Appendix 2.9.3. □

The tractable subclass of monotonic Group-Darwinian dynamics on which I focus, are those where each group responds at a constant rate. The relative rates at which groups adapt is a property I term *group-dynamism*.

**Definition 16.** Say that Group $K \in \{A, B\}$ responds at *constant-rate-*$\xi_K$, if there exists $\xi_K \in \{1, \ldots, N^K\}$, such that for all $\omega \in \Omega$, and all $t$,

$$\Psi^K(\omega_t) = [\omega_{t+1}]_K = \begin{cases} \max\{0, [\omega_t]_K - \xi_K\}, & \text{if } U^K(a;\omega_t) < U^K(b;\omega_t), \\ [\omega_t]_K, & \text{if } U^K(a;\omega_t) = U^K(b;\omega_t), \\ \min\{[\omega_t]_K + \xi_K,\, N^K\}, & \text{if } U^K(a;\omega_t) > U^K(b;\omega_t), \end{cases}$$

**Definition 17.** Say that $\Psi$ is a *constant rate dynamic* if both groups adapt at constant rates. If Groups $A$ and $B$ respond at *constant-rates* $\xi_A$ and $\xi_B$ respectively, write $\Psi = (\Psi^A_{\xi_A}, \Psi^B_{\xi_B})$.

Constant rate dynamics have a simple interpretation. Next period, a fixed number of new agents from each group adopt that group's best response (provided this new number, when added to the original number of agents who were already taking that action, does not exceed the size of the group). The *best-reply* dynamic, $\mathcal{B} := (\mathcal{B}^A, \mathcal{B}^B)$, is a constant rate dynamic with $\xi_A = N^A$ and $\xi_B = N^B$. I now define what it means for one group to be *more dynamic* than the other. Clearly, best-replying will be the most reactive a group can be.

**Definition 18.** If $\Psi$ is a constant rate dynamic, say that Group $K \in \{A, B\}$ is *(weakly) more-dynamic* than Group $K' \neq K$, written $\Psi^K \succ_d (\succeq_d)\Psi^{K'}$, if

$$\xi_{K'} < (\leq) \min\Big\{\xi_K \,,\, N^{K'}\Big\}$$

Groups adapt at the same rate, $\Psi^A \sim_d \Psi^B$, if neither is more dynamic. Formally, $\Psi^A \sim_d \Psi^B$ if either (i) $\xi_A = \xi_B = \xi \in \{1, \ldots, \min\{N^A - 1, N^B - 1\}\}$, or (ii) one of the following holds: $(\xi_A = N^A \leq \xi_B \leq N^B)$ or $(\xi_B = N^B \leq \xi_A \leq N^A)$. The first condition says that neither group best-replies, and, whenever possible, next period equal numbers of new agents from each group adopt that group's best response. The second says that if the smaller of the two groups is best-replying, then even if, whenever possible, the larger group has more agents reacting each period, both groups are still said to be evolving at the same rate.

Armed with Definitions 16-18, it will now be possible to make positive statements about varying dynamics. Normative statements are more problematic. While increased adaptiveness is a desirable property in the Language Game (Section 3.5), that is because locally risk-dominant actions coincide with most-preferred equilibrium actions. If these actions did not accord, then greater group dynamism could be detrimental.[12]

### 2.4.2 Path Dependence

In KMR, dynamics are defined on a linear state space, and when the common local-interaction is a game of pure coordination, generically there are two rest points, one at either end. KMR emphasize that *path dependence* of population behaviour rests crucially on the initial conditions, but that the final outcome is "*independent ... to all but the coarsest features of the dynamics*". The intuition for this was discussed in the story of Section 4.2. Loosely, once the process starts heading in a particular direction, it cannot "turn around". However, in the Language Game, the final outcome can depend not only on the initial state, $\omega_0$, but also on the exact specification of the dynamics.

To illustrate how path dependence may be sensitive to both the initial conditions and also the specifics of the dynamics, recall Example 1 from Section 3.3, where $\mathcal{G}^1 = \{(N^A, N^B), (p, q)\} = \{(10, 5), (3/5, 2/3)\}$. Figure 2.1 shows the state space $\Omega$ as an $11 \times 6$ lattice, with $[\omega]_A \in \{0, \ldots, 10\}$ on the horizontal-axis,

[12] That is, to be part of a less dynamic ("lazier") group may be desirable in certain situations. Thanks to Frances Ruane for originally pointing this out.

and $[\omega]_B \in \{0, \ldots, 5\}$ on the vertical-axis. Each state is depicted by a circle. The set of blue circles is $\Omega^{A,b\succeq a}$, while those red circles comprise $\Omega^{B,a\succeq b}$. These sets are defined by $(n_a^A, n_b^A, n_a^B, n_b^B) = (7, 9, 10, 6)$, calculated using equations 2.5-2.8. At the states depicted by hollow circles, $\Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$, group preferences disagree (by Lemma 1, Group $A$ prefers $a$, Group $B$ prefers $b$).

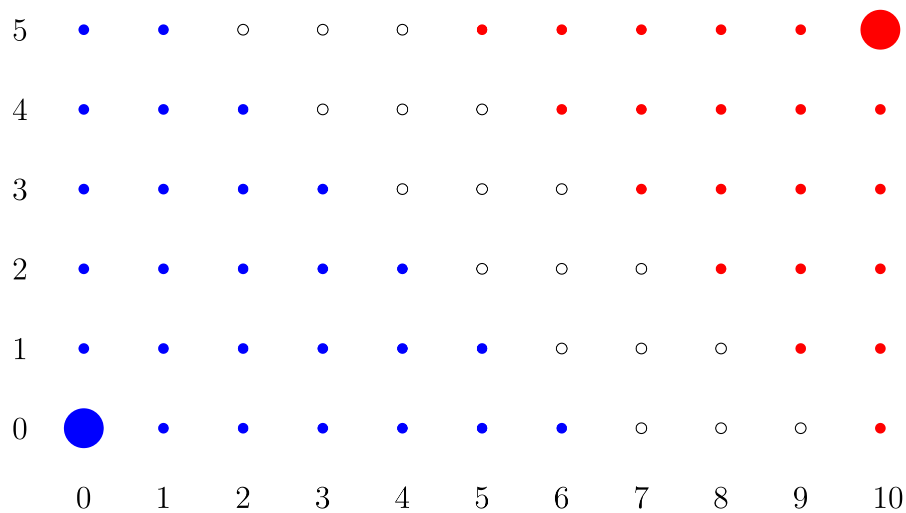


**Figure 2.1**: $\Omega^{A,b\succeq a}$ - $\Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$ - $\Omega^{B,a\succeq b}$.

A large circle denotes a rest point. Corner states $\omega_{bb}$ and $\omega_{aa}$ are always rest points, while corner state $\omega_{ab}$ is a rest point if the conditions of Theorem 3 part 4 are satisfied. While states $(n_a^A - 1, 0)$ and $(N^A, N^B - n_b^B + 1)$ can be rest points non-generically, no other state can be. Example 1 is generic and the conditions of Theorem 3 part 4 are not satisfied, and so in Figure 2.1, the only two rest points are $\omega_{bb} = (0, 0)$ and $\omega_{aa} = (10, 5)$.

To understand Group-Darwinianism, consider the state $(5, 3)$. Since $(5, 3) \in \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$, it must be that $\Psi\big((5, 3)\big) \in \{6, \ldots, 10\} \times \{0, 1, 2\}$. While monotonicity is not shown in Figure 2.1, it is also easily understood when coupled with Group-Darwinianism. Consider the pair of states, $(5, 2)$ and $(5, 3)$. Clearly $(5, 3) \geqslant_a (5, 2)$, and so $\Psi\big((5, 3)\big) \geqslant_a \Psi\big((5, 2)\big)$. Since $(5, 2) \in \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$, $\Psi\big((5, 2)\big) \in \{(6, \ldots 10\} \times \{0, 1\}$. Now suppose that $\Psi\big((5, 2)\big) = (8, 1)$. Then $\Psi\big((5, 3)\big)$ is further restricted to the set $\Big\{(8, 1), (9, 1), (10, 1), (8, 2), (9, 2), (10, 2)\Big\}$.

Figure 2.1 is a preference map, with $\left\{\Omega^{A,b\succeq a}, \Omega^{B,a\succeq b}, \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}\right\}$ as the partition. To partition $\Omega$ into $\{\mathcal{V}(\omega)\}_{\omega\in\Omega_0}$, further information on the details of the dynamics are needed. To see how these details can matter, it is instructive to start by looking at behaviour at individual states for varying dynamics.[13]

**Example 3** ($\Psi = (\mathcal{B}^A, \mathcal{B}^B)$)**.** From Figure 2.1, the interpretation is easy: no matter what the current state, the following state must be $\omega_{bb}, \omega_{ab}$, or $\omega_{aa}$. Blue states jump immediately to $\omega_{bb}$, black states to $\omega_{ab}$, and red states to $\omega_{aa}$. Formally, for any $\omega \in \Omega^{A,b\succeq a}$, $\Psi(\omega) = \mathcal{B}(\omega) = \omega_{bb}$; for any $\omega \in \Omega^{B,a\succeq b}$, $\Psi(\omega) = \mathcal{B}(\omega) = \omega_{aa}$; and for any $\omega \in \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$, $\mathcal{B}(\omega) = \omega_{ab}$ and $\mathcal{B}^2(\omega) = \omega_{aa}$. That is, states in $\Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$ transition first to $\omega_{ab}$, and from there to $\omega_{aa}$, and so with $\Psi = \mathcal{B}$, are considered red.

**Example 4** ($\Psi = (\Psi_1^A, \mathcal{B}^B)$)**.** Figure 2.2 illustrates the basins of attraction for this scenario. As before, states in $\Omega^{B,a\succeq b}$ are denoted by solid red circles, and states in $\Omega^{A,b\succeq a}$ by solid blue circles. States for which the groups have conflicting preferences, $\Omega^{B,b\succ a} \cap \Omega^{A,a\succ b}$, are again denoted by hollow circles. However, it is not the case that $\Omega^{B,b\succ a} \cap \Omega^{A,a\succ b} \subseteq \mathcal{V}(\omega_{aa})$, as it was when $\Psi = \mathcal{B}$. Hollow red circles eventually lead to $\Omega^{B,a\succeq b}$ and hence to $\omega_{aa}$, but hollow blue circles lead immediately to $\Omega^{A,b\succeq a}$ and eventually to $\omega_{bb}$.

To be clear what is happening, again consider state $(5,3)$. When $\Psi = (\mathcal{B}^A, \mathcal{B}^B)$ as in Example 3, the dynamics terminate at $(10,5)$ via the path $\{(5,3) \to (10,0) \to (10,5)\}$. When $\Psi = (\Psi_1^A, \mathcal{B}^B)$ as in Example 4, $(5,3)$ leads to $(0,0)$, via the path $\{(5,3) \to (6,0) \to (5,0) \to (4,0) \to (3,0) \to (2,0) \to (1,0) \to (0,0)\}$.

We can also trace how the boundaries of the basins of attraction vary.[14] When $\Psi^A \sim_d \Psi^B$, all basins of attraction, and hence their boundaries, are the same.[15] By Lemma 3, for any monotone dynamics the basins will be convex. This is easily seen by inspection of Figures 2.1 and 2.2. From Figure 2.1, and for any $\Psi$ such that $\xi_A = \xi_B$,

[13] The formal analysis of how basins of attraction vary across constant rate adaptive processes is carried out in Appendix 2.9.2.

[14] The boundaries, both upper (lower) and total upper (total lower), are defined in Appendix 2.9.1.

[15] In fact, this statement holds true under the weaker condition that $\xi_A \geq n_a^A$ and $\xi_B \geq n_b^B$.

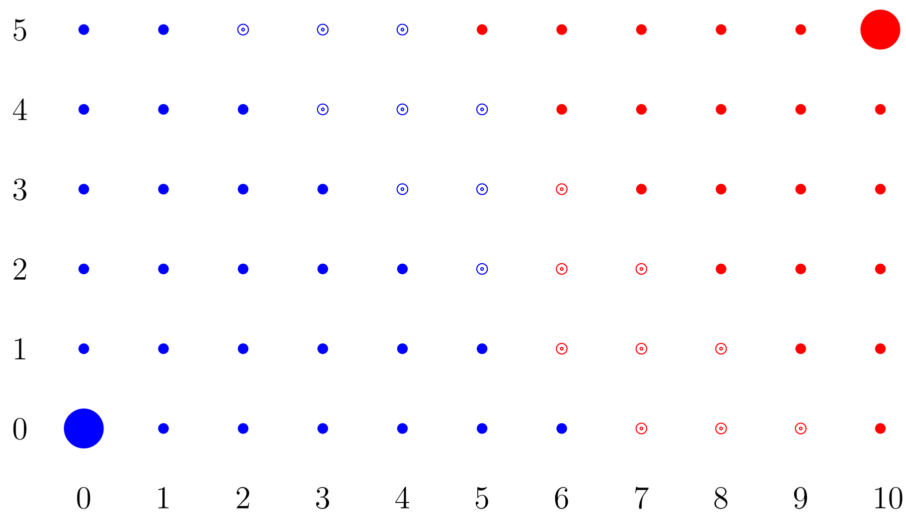


**Figure 2.2**: $\mathcal{V}(\omega_{bb})$ and $\mathcal{V}(\omega_{aa})$ when $\Psi = (\Psi_1^A, \mathcal{B}^B)$.

- $\big(\mathcal{V}(\omega_{bb})\big)_+ = \big(\mathcal{V}(\omega_{bb})\big)_{++} = \{(1,5),(2,4),(3,3),(4,2),(5,1),(6,0)\}$
- $\big(\mathcal{V}(\omega_{aa})\big)_- = \big(\mathcal{V}(\omega_{aa})\big)_{--} = \{(2,5),(3,4),(4,3),(5,2),(6,1),(7,0)\}$

while in Figure 2.2, with $\Psi = (\Psi_1^A, \mathcal{B}^B)$,

- $\big(\mathcal{V}(\omega_{bb})\big)_+ = \{(4,5),(5,4),(6,0)\}$
- $\big(\mathcal{V}(\omega_{bb})\big)_{++} = \{(4,5),(5,4),(5,3),(5,2),(5,1),(6,0)\}$
- $\big(\mathcal{V}(\omega_{aa})\big)_- = \{(5,5),(6,1),(7,0)\}$
- $\big(\mathcal{V}(\omega_{aa})\big)_{--} = \{(5,5),(6,4),(6,3),(6,2),(6,1),(7,0)\}$

## 2.5 Equilibrium Selection

Any deterministic Group-Darwinian dynamic, $\Psi$, induces a time homogeneous Markov process on the finite state space $\Omega$. Let $P$ be the associated Markov matrix, where for every pair of states $\omega', \omega'' \in \Omega$, $P(\omega', \omega'') \geq 0$ denotes the probability of transitioning from $\omega'$ to $\omega''$, and for each $\omega \in \Omega$, $\sum_{\omega'} P(\omega, \omega') = 1$.

For any finite set $X$, let $\triangle(X)$ denote the set of distributions on $X$. A stationary distribution of $P$ is a row-vector $\mu \in \triangle(\Omega)$, such that $\mu P = \mu$. The set of stationary distributions is denoted $\triangle_0(\Omega)$. Writing supp($\mu$) for the support of $\mu$, say that $D \subset \Omega$ is a recurrent class, if for all $\omega \in \Omega$, and all $\mu \in \triangle_0(\Omega)$ with supp($\mu$) $\subset D$, $\mu(\omega) > 0 \iff \omega \in D$. A state is *recurrent* if it is contained in a recurrent class, and *transient* otherwise. A singleton recurrent class is an *absorbing state*.

All Markov processes possess at least one stationary distribution, while *ergodic* Markov processes possess only one. The third assumption of KMR, perturbs the deterministic dynamics in such a way as to induce a new Markov process that is ergodic.[16]

**Assumption 3.** *Behavioural Mutation: There is a small probability that an agent may choose an action at random.*

One interpretation is as follows. After afforded decisions have been taken, but before payoffs are made, with probability $\varepsilon^A > 0$ ($\varepsilon^B > 0$) each Group $A$ ($B$) player switches his current action choice, and with probability $1 - \varepsilon^A$ ($1 - \varepsilon^B$) maintains his action.[17]

Even for constant rate dynamics, there is no grounds for always assuming $\Psi^A \sim_d \Psi^B$. Similarly, there is no reason to suppose that behavioural mutations occur with equal likelihood for members of different groups.[18] So while interest will lie in the case where $(\varepsilon^A, \varepsilon^B) \to (0, 0)$, I will be making the strong assumption that $\varepsilon^A = \varepsilon^B = \varepsilon$, for all states and all time periods. It is tempting to insist on a milder condition like $\varepsilon^A = O(\varepsilon^B)$ and $\varepsilon^B = O(\varepsilon^A)$,[19] but the selection results may differ.[20]

[16]Technically this assumption makes the process *irreducible* and *aperiodic* which for finite state Markov processes is sufficient for ergodicity. See Karlin and Taylor (1975).

[17]This is a different interpretation to that given in KMR, but it generates the same switching probabilities.

[18]Nor is there any reason to suppose mutations are both state- and time-independent. The effects that subtle differences in mutation rates can have on equilibrium selection are examined in Bergin and Lipman (1996).

[19]Letting both $\varepsilon^A$ and $\varepsilon^B$ be functions of $\varepsilon$, we say that $\varepsilon^A(\varepsilon) = O\big(\varepsilon^B(\varepsilon)\big)$ as $\varepsilon \downarrow 0$, if and only if there exists positive numbers $M$ and $\delta$, such that $\varepsilon^A(\varepsilon) \leq |M\varepsilon^B(\varepsilon)|$ for all $\varepsilon < \delta$.

[20]However, this should not be muddled with the results of Bergin and Lipman (1996) (see footnote 4 above). To my knowledge, there are no games for which the Ising model dynamics of Blume (1993) select *different* long run equilibria to the dynamics of KMR. In Neary (2010b), I show how these two dynamics can arrive at different selection results for an open set of parameters for the Language Game.

For a given $\varepsilon > 0$, the above perturbations define a new ergodic Markov process with associated transition matrix $P^\varepsilon$, and unique stationary distribution $\mu^\varepsilon$. By continuity, the accumulation point of $\{\mu^\varepsilon\}_{\varepsilon>0}$, $\mu^\star$, is a stationary distribution of $P := \lim_{\varepsilon\downarrow 0} P^\varepsilon$. Our interest lies in the states to which $\mu^\star$ assigns positive probability.

**Definition 19.** State $\omega$ is *stochastically stable* if $\mu^\star(\omega) > 0$, and *uniquely stochastically stable* if $\mu^\star(\omega) = 1$. Let $\Omega^\star$ denote the set of stochastically stable states.

Write $\mathcal{L}$ for the collection of recurrent classes. For the Language Game, it can be shown that $\mathcal{L} = \{\{\omega\}\}_{\omega\in\Omega_0}$, where $\Omega_0$ is the set of rest points as defined in Section 3.4. We can express $\Omega^\star$ as the union of recurrent classes as follows

$$\exists\, \mathcal{M} \subset \mathcal{L} \text{ such that } \Omega^\star = \bigcup_{D\in\mathcal{M}} D$$

Calculating $\mu^\star$ is the objective. This is done using tree-surgery techniques from Freidlin and Wentzell (1998), first introduced to game theory in Foster and Young (1990). To do so, it will be useful to view states in $\Omega$ as the vertices of a fully connected directed graph, $\Gamma^\star$. An edge in $\Gamma^\star$ from $\omega'$ to $\omega''$ is denoted $(\omega' \to \omega'')$. A *walk* from $\omega'$ to $\omega''$ is a sequence of edges $\{(\omega_i \to \omega_{i+1})\}_{i=0}^{m-1}$ where $\omega_0 = \omega'$, and $\omega_m = \omega''$. A path is a walk in which the vertices are distinct. A typical path from $\omega'$ to $\omega''$ is denoted by $h(\omega', \omega'')$, and the set of all paths from $\omega'$ to $\omega''$ by $H(\omega', \omega'')$. Extending this, the set of all paths from a state $\omega$ to a set $Q \not\ni \omega$ can be defined as follows, $H(\omega, Q) = \cup_{\omega'\in Q} H(\omega, \omega')$.

Following KMR, we assume that for any pair $\omega', \omega''$, the following limit exists

$$\begin{aligned} c_\Psi(\omega', \omega'') &:= \lim_{\varepsilon\downarrow 0} \frac{\log P^\varepsilon(\omega', \omega'')}{\log \varepsilon} \\ &= \|\Psi(\omega'), \omega''\| \end{aligned}$$

where $c_\Psi : \Omega \times \Omega \to \mathbb{R}_+ \cup \{\infty\}$ is a *cost function*. The value it takes for a particular pair $(\omega', \omega'')$ is interpreted as the minimum number of simultaneous mutations needed to transition directly from $\omega'$ to $\omega''$, or in graph theoretic terms, as the cost of edge $(\omega' \to \omega'')$ in $\Gamma^\star$.[21]

For any function $\tau : \Omega \to \Omega$, a path from $\omega'$ to $\omega''$ in $\tau$, is a path $\{(\omega_0 \to \omega_1), (\omega_1 \to \omega_2), \ldots, (\omega_{m-1} \to \omega_m)\}$, where $\omega_0 = \omega'$ and $\omega_m = \omega''$, such that $\tau(\omega_i) = \omega_{i+1}$ for all $i = 0, \ldots, m-1$. An $\omega$-tree, $\tau_\omega$, is a mapping $\tau_\omega : \Omega \to \Omega$ such that: (i) $\tau_\omega(\omega) = \omega$; (ii) for every $\omega' \in \Omega \setminus \{\omega\}$, there is a unique path in $\tau_\omega$ from $\omega'$ to $\omega$. Say that $\omega''$ is a *successor* of $\omega'$ in $\tau_\omega$ if $\tau_\omega^m(\omega') = \omega''$ for some $m \geq 1$, and the *immediate successor* if $m = 1$.

For each $\omega$, $T_\omega$ is the set of all $\omega$-trees. The cost of $\omega$-tree, $\tau_\omega \in T_\omega$, is the sum of the costs of its edges,

$$c_\Psi(\tau_\omega) = \sum_{\omega' \neq \omega} c_\Psi\big(\omega', \tau_\omega(\omega')\big)$$

For the Language Game, $\mathcal{G}$, and cost function, $c_\Psi$, define the set of states that achieve minimum cost $\omega$-trees as

$$\Xi(\mathcal{G}, c_\Psi) := \left\{ \omega^\star \in \Omega \;\middle|\; \text{for any } \omega \in \Omega, \min_{\tau_{\omega^\star} \in T_{\omega^\star}} c_\Psi(\tau_{\omega^\star}) \leq \min_{\tau_\omega \in T_\omega} c_\Psi(\tau_\omega) \right\}$$

The following is the result of Freidlin and Wentzell (1998). Note it's relation to Definition 19 above.

**Lemma 4.** *State $\omega$ is stochastically stable if $\omega \in \Xi(\mathcal{G}, c_\Psi)$, and uniquely stochastically stable if $\{\omega\} = \Xi(\mathcal{G}, c_\Psi)$. That is, $\Xi(\mathcal{G}, c_\Psi) = \Omega^\star$.*

By Young (1993), Theorem 4, the stochastically stable states are contained in a recurrent class. We can therefore restrict attention to minimum cost $\omega$-trees of recurrent states. For the Language Game, $\mathcal{L} = \{\{\omega\}\}_{\omega \in \Omega_0}$, and $\mathcal{V}(\omega') \cap \mathcal{V}(\omega'') = \emptyset$ for all distinct $\omega', \omega'' \in \Omega_0$. Thus, the key to computing $\omega$-trees of the absorbing states, is to find a path of minimum cost from each absorbing state to the convex basin of attraction of the others. For any pair of states $\omega', \omega'' \in \Omega$, denote by

[21] Note that $c_\Psi$ is allowed to take the value $\infty$. This could be the case if a transition is impossible under the dynamics. In this paper, the range of $c_\Psi$ is the range of $\|\,,\|$ which is $\{0, \ldots, N\}$.

$c^{\star}_{\Psi}(\omega', \omega'')$ the cost of the path of minimum cost between them. That is,

$$c^{\star}_{\Psi}(\omega', \omega'') := \min_{\{(\omega_j \to \omega_{j+1})\}_{j=0}^{n-1} \in H(\omega', \omega'')} \sum_{m=0}^{n-1} c_{\Psi}(\omega_m, \omega_{m+1})$$

The two main results in this Section concern equilibrium selection. Before presenting these however, the following Lemma is needed. It says that a path of minimum cost out of a region of the state space in which the dynamics are unambiguous, involves a direct transition out. An immediate and important consequence is that it holds for the symmetric profiles $\omega_{bb}$ and $\omega_{aa}$.

**Lemma 5.** *Let $\Psi$ be a monotonic Group-Darwinian adjustment process. Then,*

1. *For all $\omega \in \Omega^{A,b\succeq a}$, the minimum of $c_{\Psi}(h')$ over all paths $h' \in H\big(\omega, \Omega\backslash\Omega^{A,b\succeq a}\big)$ is attained by*

$$h^{\star} := \Big\{(\omega \to \omega^{\star})\Big\}$$

*where*

$$\omega^{\star} \in \operatorname*{argmin}_{\hat{\omega} \,\in\, (\Omega\backslash\Omega^{A,b\succeq a})_{-}} \|\Psi(\omega), \hat{\omega}\|$$

2. *For all $\omega \in \Omega^{B,a\succeq b}$, the minimum of $c_{\Psi}(h'')$ over all paths $h'' \in H\big(\omega, \Omega\backslash\Omega^{B,a\succeq b}\big)$ is attained by*

$$h^{\star\star} := \Big\{(\omega \to \omega^{\star\star})\Big\}$$

*where*

$$\omega^{\star\star} \in \operatorname*{argmin}_{\hat{\omega} \,\in\, (\Omega\backslash\Omega^{B,a\succeq b})_{+}} \|\Psi(\omega), \hat{\omega}\|$$

*However, the paths $h^{\star}$ and $h^{\star\star}$ above attaining these minimum costs need not be unique.*

*Proof.* The proof is found in Appendix 2.9.3. □

Generically $\Omega^{A,b\succeq a} \subseteq \mathcal{V}(\omega_{bb})$ and $\Omega^{B,a\succeq b} \subseteq \mathcal{V}(\omega_{aa})$. So all that remains is to classify behaviour for states in $\big(\Omega^{A,a\succ b} \cap \Omega^{B,a\succ b}\big)$. It turns out that the behaviour of the dynamics in these states can be key for selection. When $E(\mathcal{G}) = \{\omega_{bb}, \omega_{aa}\}$, the set $\left(\Omega^{A,a\succ b} \cap \Omega^{B,a\succ b}\right)$ has more of an effect than when $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$. The analysis of each case is quite different and so are looked at separately. The second case is easier to begin with.

### 2.5.1 Equilibrium set is $\{\omega_{\mathbf{bb}}, \omega_{\mathbf{ab}}, \omega_{\mathbf{aa}}\}$

We begin by calculating the minimum cost $\omega$-trees of each convention, when the groups adapt at constant and equal rates.

**Theorem 4.** *Suppose Condition 4 of Theorem 3 holds, so $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, and that the monotonic Group-Darwinian adjustment process is such that both groups evolve at constant and equal rates. Let $\tau^\star_{\omega_{bb}}$, $\tau^\star_{\omega_{ab}}$, and $\tau^\star_{\omega_{aa}}$, denote minimum cost $\omega$-trees for $\omega_{bb}, \omega_{ab}$, and $\omega_{aa}$ respectively. Then,*

$$c_\Psi(\tau^\star_{\omega_{bb}}) = n_b^B + n_b^A - N^B \tag{2.15}$$

$$c_\Psi(\tau^\star_{\omega_{ab}}) = n_b^B + n_a^A \tag{2.16}$$

$$c_\Psi(\tau^\star_{\omega_{aa}}) = n_a^A + n_a^B - N^A \tag{2.17}$$

*The set of stochastically stable states are those with $\omega$-tree of minimum cost. That is,*

$$\Xi(\mathcal{G}, c_\Psi) = \underset{\omega \in \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}}{\operatorname{argmin}} c_\Psi(\tau^\star_\omega)$$

The details are found in Appendix 2.9.3. Here I discuss the intuition. The proof rests on computing paths of minimum cost between the six pairs of coventions, $(\omega_{bb}, \omega_{ab}), (\omega_{bb}, \omega_{aa}), (\omega_{ab}, \omega_{bb}), (\omega_{ab}, \omega_{aa}), (\omega_{aa}, \omega_{bb})$, and $(\omega_{aa}, \omega_{ab})$. By Lemma 5, the minima of $c_\Psi(h')$ over all $h' \in H(\omega_{bb}, \omega_{ab})$, and of $c_\Psi(h'')$ over all $h'' \in H(\omega_{aa}, \omega_{ab})$, are attained by

$$\begin{aligned}
h^\star(\omega_{bb}, \omega_{ab}) = &\left\{\left(\omega_{bb} \to (n_a^A, 0)\right)\right\} \\
&\cup \left\{\left(\omega' \to \Psi(\omega')\right) \mid \omega' = \Psi^m\left((n_a^A, 0)\right) \text{ for some } m \geq 0\right\} \\
h^\star(\omega_{aa}, \omega_{ab}) = &\left\{\left(\omega_{aa} \to (N^A, N^B - n_b^B)\right)\right\} \\
&\cup \left\{\left(\omega' \to \Psi(\omega')\right) \mid \omega' = \Psi^m\left((N^A, N^B - n_b^B)\right) \text{ for some } m \geq 0\right\}
\end{aligned}$$

Next, use Lemma 8 to show that the minima of $c_\Psi(h')$ over all $h' \in H(\omega_{ab}, \omega_{bb})$, and of $c_\Psi(h'')$ over all $h'' \in H(\omega_{ab}, \omega_{aa})$, are attained by

$$\begin{aligned} h^\star(\omega_{ab}, \omega_{bb}) = &\big\{\big(\omega_{ab} \to (n_b^A - N^B, 0)\big)\big\} \\ &\cup \Big\{\big(\omega' \to \Psi(\omega')\big) \,|\, \omega' = \Psi^m\big((n_b^A - N^B, 0)\big) \text{ for some } m \geq 0\Big\} \\ h^\star(\omega_{ab}, \omega_{aa}) = &\big\{\big(\omega_{ab} \to (N^A, n_a^B - N^A)\big)\big\} \\ &\cup \Big\{\big(\omega' \to \Psi(\omega')\big) \,|\, \omega' = \Psi^m\big((N^A, n_a^B - N^A)\big) \text{ for some } m \geq 0\Big\} \end{aligned}$$

Lastly, note that when both groups adapt at constant and equal rates, $\mathcal{V}(\omega_{ab})$ is "sandwiched" between $\mathcal{V}(\omega_{bb})$ and $\mathcal{V}(\omega_{aa})$. That is, for all $\omega' \in \mathcal{V}(\omega_{bb})$ and $\omega'' \in \mathcal{V}(\omega_{aa})$ with $\omega' \leqslant_a \omega''$, there exists $\hat{\omega} \in \mathcal{V}(\omega_{ab})$ such that $\omega' \leqslant_a \hat{\omega} \leqslant_a \omega''$. Using this, it is easily shown that paths of minimum cost from $\omega_{bb}$ to $\omega_{aa}$, and from $\omega_{aa}$ to $\omega_{bb}$ are given by

$$\begin{aligned} h^\star(\omega_{bb}, \omega_{aa}) &= h^\star(\omega_{bb}, \omega_{ab}) \cup h^\star(\omega_{ab}, \omega_{aa}) \\ h^\star(\omega_{aa}, \omega_{bb}) &= h^\star(\omega_{aa}, \omega_{ab}) \cup h^\star(\omega_{ab}, \omega_{bb}) \end{aligned}$$

where $h^\star(\omega_{bb}, \omega_{ab}), h^\star(\omega_{ab}, \omega_{aa}), h^\star(\omega_{aa}, \omega_{ab})$, and $h^\star(\omega_{ab}, \omega_{bb})$ are as given above. That these are the only costly paths of the respective $\omega$-trees with costs as given in equations 2.15-2.17 is clear. That the set of stochastically stable states are those with $\omega$-tree of minimum cost is immediate by Lemma 4.

I now show (Theorem 5) that when $\omega_{ab}$ is stochastically stable under constant rate dynamics where both groups adapt at equal rates, the set of stochastically stable equilibria is independent of the specifics of the dynamics. What happens is this. First, varying rates of adjustment will never lower the cost $\tau^\star_{\omega_{ab}}$. Showing this is straightforward. Second, while it may lower the cost of $\tau^\star_{\omega_{bb}}$ or $\tau^\star_{\omega_{aa}}$, it will not lower the cost enough to alter selection. That is, Theorem 5 does not say that the minimum cost $\omega$-tree of each convention is necessarily unchanged and as given by equations 2.15-2.17. Rather it just says that if $\omega_{ab}$ is ever stochastically stable for some constant rate dynamic, it will always be for any constant rate dynamic and it will have $\omega$-tree with cost given by that in Theorem 4.

**Theorem 5.** *Suppose Condition 4 of Theorem 3 holds, so* $E = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$*, and that the monotonic Group-Darwinian adjustment process is such that both groups evolve at constant rates. If* $\omega_{ab} \in \Xi(\mathcal{G}, c_\Psi)$ *when* $\Psi^A \sim_d \Psi^B$*, then* $\omega_{ab} \in \Xi(\mathcal{G}, c_\Psi)$ *for any constant rate adaptive process. Futhermore, states in* $\Xi(\mathcal{G}, c_\Psi)$ *have* $\omega$*-tree of minimum cost equal to that as given in Theorem 4.*

*Proof.* The proof is found in Appendix 2.9.3. □

This is a good time to mention a few features of the set up. Ellison (2000) introduced the notions of the *radius* and *coradius* of a recurrent class. The radius is defined as the minimum number of mutations necessary to escape the basin, while the coradius is defined as the maximum (over all states) of the minimum number of mutations necessary to reach the basin. When the radius is greater than the coradius, then the long run equilibrium belongs to this recurrent class. The result is not universally powerful since it does not apply in all cases. In the Language Game, it need not have relevance when $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, since it is possible that each rest point's coradius is larger than its radius.

Another observation is that due to the 2-dimensional nature of the state space, there need not be a connection between the size of each convention's basin and stochastic stability.[22] In fact, it is very possible that the equilibrium with the largest basin of attraction is not stochastically stable, and that the equilibrium with the smallest basin is stochastically stable. This can be particularly striking for parameters for which $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$.

The following example illustrates both these phenomena, and also provides the intuition behind Theorems 4 and 5. For completeness' sake, we choose non-generic parameters.

**Example 5.** Let $\mathcal{G}^3 = (10, 10, 4/5, 2/3)$.
$E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$ for these parameters. Using equations 2.5 - 2.8, get $(n_a^A, n_b^A, n_a^B, n_b^B) = (5, 16, 7, 13)$.

Figure 2.3 below illustrates the basins of attraction when both groups respond at equal rates. The non-genericity is on display by denoting the non-

[22] A similar though slightly different point is made in Section 8 of Young (1993).

equilibrium rest point $(10,3)$ by an X. The sizes of the basins of attraction are $|\mathcal{V}\big((10,3)\big)| = 1$, $|\mathcal{V}(\omega_{bb})| = 28$, $|\mathcal{V}(\omega_{ab})| = 57$, and $|\mathcal{V}(\omega_{aa})| = 35$. Clearly $\mathcal{V}(\omega_{ab})$ is the largest.

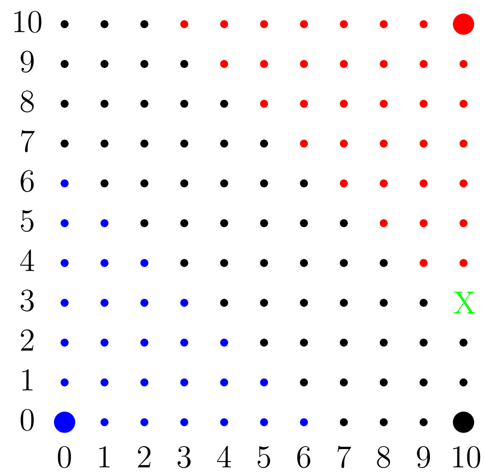


**Figure 2.3**: $\mathcal{V}(\omega_{bb}) = \bullet$. $\mathcal{V}(\omega_{ab}) = \bullet$. $\mathcal{V}(\omega_{aa}) = \bullet$.

Edges of positive cost in $\tau^{\star}_{\omega_{aa}}$ are $\big((0,0) \to (7,0)\big)$, $\big((10,0) \to (10,3)\big)$, and $\big((10,3) \to (10,4)\big)$. These have a combined cost of $7+3+1 = 11$. The costs of $\tau^{\star}_{(10,3)}$, $\tau^{\star}_{\omega_{bb}}$, and $\tau^{\star}_{\omega_{ab}}$ can be computed as 17, 12, and 15 respectively. By Theorem 4, $\omega_{aa}$ is the selected long run equilibrium, and this is despite it not having the maximal basin of attraction.

The [radius, coradius] pair for each absorbing state can also easily be calculated. They are given as follows, $\omega_{bb} \mapsto [7,12]$, $\omega_{ab} \mapsto [3,8]$, $\omega_{aa} \mapsto [7,11]$, and $(10,3) \mapsto [1,10]$. Conventions $\omega_{bb}$ and $\omega_{aa}$ have equal radii, while $\omega_{bb}$ has maximal coradius. Note that all rest points have a greater coradius than radius so the Theorem of Ellison (2000) does not apply.

Let us now vary the rates of reaction for each group. First of all, note that regardless of rates, there exist some states at which the dynamics are unambiguous. These states are illustrated in Figure 2.4 and are colour coded by the convention to which they lead. The states not shown are a subset of $\Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$, are were part of $\mathcal{V}(\omega_{ab})$ in Figure 2.3 when both groups responded at equal rates. We will demonstrate how Figure 2.4 is modified when the groups adapt at different rates.

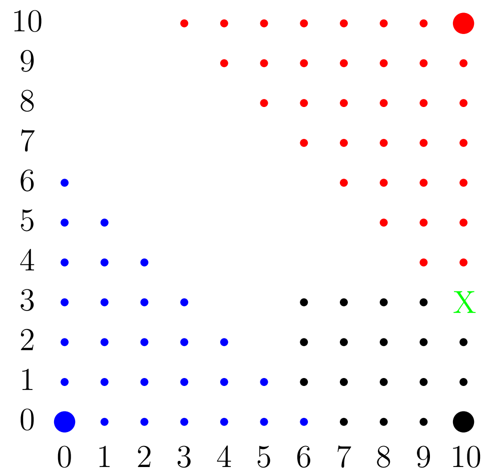


**Figure 2.4**: $\mathcal{V}(\omega_{bb}) = \bullet$. $\mathcal{V}(\omega_{ab}) = \bullet$. $\mathcal{V}(\omega_{aa}) = \bullet$.

Begin by supposing $\Psi = (\mathcal{B}^A, \Psi_1^B)$. Basins of attraction for this case are illustrated in Figure 2.5. Rest point $(10, 3)$ now has a non-degenerate basin of attraction, with states denoted by $\times$s. Basin $\mathcal{V}(\omega_{bb})$ is as it was when $\Psi^A \sim_d \Psi^B$. It is now the case that $\mathcal{V}(\omega_{aa}) = \Omega^{B,a\succeq b} \cup \{\omega \in \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a} \mid [\omega]_B \geq 5\}$. The sizes of the basins have changed. Now, $|\mathcal{V}((10,3))| = 7$, $|\mathcal{V}(\omega_{bb})| = 28$, $|\mathcal{V}(\omega_{ab})| = 21$, and $|\mathcal{V}(\omega_{aa})| = 65$. By Theorem 5, it must still be that $\Xi(\mathcal{G}, c_\Psi) = \{\omega_{aa}\}$. It is easily calculated that minimum cost $\omega$-trees for the absorbing states have not changed.

While radii are always unaffected by varying reaction rates, the coradii of absorbing states $\omega_{aa}$ and $(10, 3)$ have changed. The [radius, coradius] pairs are now given as follows, $\omega_{bb} \mapsto [7, 12]$, $\omega_{ab} \mapsto [3, 8]$, $\omega_{aa} \mapsto [7, 7]$, and $(10, 3) \mapsto [1, 7]$. It is still the case that no rest point has a greater coradius than radius so the result of Ellison (2000) remains inapplicable.

Now suppose $\Psi = (\Psi_1^A, \mathcal{B}^B)$. Basins of attraction are illustrated in Figure 2.6. We now have $|\mathcal{V}((10,3))| = 1$, $|\mathcal{V}(\omega_{bb})| = 61$, $|\mathcal{V}(\omega_{ab})| = 24$, and $|\mathcal{V}(\omega_{aa})| = 35$. By Theorem 4, the minimum cost $\omega_{aa}$-, $\omega_{ab}$-, and $(10, 3)$-trees still have cost of 11, 15, and 17 respectively. Consider the minimum cost $\omega_{bb}$-tree. The cost of 12 remains an upper bound. But an $\omega_{bb}$-tree with cost 12 is also attainable using costly paths $((10, 0) \to (10, 3))$, $((10, 3) \to (10, 4))$, and $((10, 10) \to (2, 10))$. This

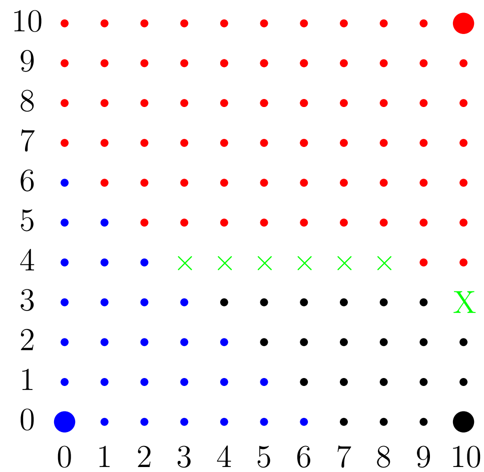

**Figure 2.5**: $\mathcal{V}(\omega_{bb}) = \bullet$. $\mathcal{V}(\omega_{ab}) = \bullet$. $\mathcal{V}(\omega_{aa}) = \bullet$.

$\omega_{bb}$-tree also has a total cost of 12.

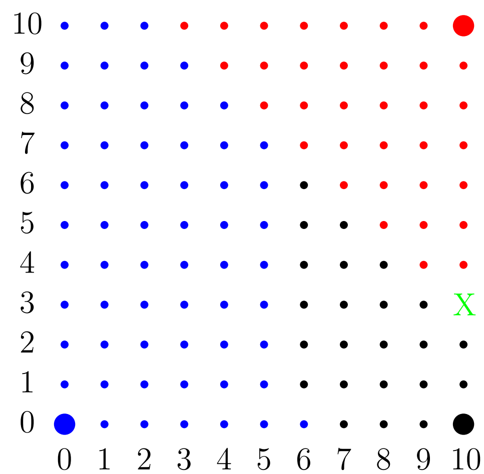

**Figure 2.6**: $\mathcal{V}(\omega_{bb}) = \bullet$. $\mathcal{V}(\omega_{ab}) = \bullet$. $\mathcal{V}(\omega_{aa}) = \bullet$.

The following modification hints at how rates can affect minimum cost $\omega$-trees enough to alter selection. Suppose for a moment that the Group $B$ payoffs were perturbed so that the game became generic with $(10, 3) \in \Omega^{A,b \succ a}$. With $k_A = k_B$, we get $c_\Psi(\tau^\star_{\omega_{bb}}) = 12$, $c_\Psi(\tau^\star_{\omega_{ab}}) = 15$, and $c_\Psi(\tau^\star_{\omega_{aa}}) = 10$. Suppose now that $\Psi = (\Psi^A_1, \mathcal{B}^B)$. The minimum cost $\omega_{aa}$-tree still has a cost 10. The cost of the minimum cost $\omega_{bb}$-tree however, has lowered to 11 by proceeding along the route described in the previous paragraph. But lowering it from 12 to 11 is not enough to affect equilibrium selection in this case. However, this feature is not robust as was stated in Theorem 5.

### 2.5.2 Equilibrium set is $\{\omega_{\mathbf{bb}}, \omega_{\mathbf{aa}}\}$

Now we examine the case when only the symmetric profiles are equilibria.

**Theorem 6.** *Suppose Condition 4 of Theorem 3 does not hold, so* $E = \{\omega_{bb}, \omega_{aa}\}$, *and that the monotonic Group-Darwinian adjustment process is such that both groups evolve at constant rates,* $\xi_A$ *and* $\xi_B$ *respectively. Let* $\tau^\star_{\omega_{bb}}$ *and* $\tau^\star_{\omega_{aa}}$, *denote the minimum cost* $\omega$*-trees for* $\omega_{bb}$ *and* $\omega_{aa}$ *respectively. Then,*

1. *If* $(N^A, 0) \in \mathcal{V}(\omega_{aa})$, *then*

$$\begin{aligned}
c_\Psi(\tau^\star_{\omega_{aa}}) &= n^A_a \\
c_\Psi(\tau^\star_{\omega_{bb}}) &= n^A_b - 1_{\{\Psi^B \succ_d \Psi^A\}}\left[\max_{\omega'' \in (\mathcal{V}(\omega_{bb}))_+} \; \min_{\omega' \in (\Omega^{B,a\succeq b})_+} \|\omega', \omega''\|\right] \\
&= n^A_b - \left\| \big((\mathcal{V}(\omega_{bb}))_+\big)_{NW}, \; \big((\Omega^{A,b\succeq a})_+\big)_{NW} \right\|
\end{aligned}$$

2. *If* $(N^A, 0) \in \mathcal{V}(\omega_{bb})$, *then*

$$\begin{aligned}
c_\Psi(\tau^\star_{\omega_{bb}}) &= n^B_b \\
c_\Psi(\tau^\star_{\omega_{aa}}) &= n^B_a - 1_{\{\Psi^A \succ_d \Psi^B\}}\left[\max_{\omega'' \in (\mathcal{V}(\omega_{aa}))_-} \; \min_{\omega' \in (\Omega^{A,b\succeq a})_-} \|\omega', \omega''\|\right] \\
&= n^B_a - \left\| \big((\mathcal{V}(\omega_{aa}))_-\big)_{NW}, \; (\Omega^{A,b\succeq a})_- \right\|
\end{aligned}$$

*The set of stochastically stable states are those with the* $\omega$*-tree of minimum cost. That is,*

$$\Xi(\mathcal{G}, c_\Psi) = \operatorname*{argmin}_{\omega \in \{\omega_{bb}, \omega_{aa}\}} c_\Psi(\tau^\star_\omega)$$

*Proof.* The proof is found in Appendix 2.9.3. □

Some things are worth noting. Clearly the state $\omega_{ab}$ plays an important role. Since it is not an equilibrium, it always lies in the basin of attraction of the larger group's preferred convention. If $N^A > N^B$ and $\Psi^A \succ_d \Psi^B$, the set $\Omega\backslash(\Omega^{A,b\succeq a} \cup \Omega^{B,a\succeq b})$ is a subset of $\mathcal{V}(\omega_{aa})$. When $\omega_{ab} \in \mathcal{V}(\omega_{aa})$, then the selected long run equilibrium might change if $\Psi^B \succ_d \Psi^A$. When $\omega_{ab} \in \mathcal{V}(\omega_{bb})$ the reverse holds. Appendix 2.9.2 shows how to construct basins of attraction when the smaller group responds at a faster rate.

Suppose Group $A$ is the larger group. The minimum cost of an $\omega_{aa}$-tree is $n_a^A$ (since $(n_a^A, 0) \in \mathcal{V}(\omega_{aa})$), while a lower bound for the minimum cost $\omega_{bb}$-tree is $n_b^B$ (if $(N^A - n_b^B, N^B)$). Thus, Group $B$ must have a stronger relative preference for its preferred equilibrium to reverse the long run outcome. Conversely, if Group $A$ has an equally strong preference, then $(\mathbf{a}, \mathbf{a})$ will always be a stochastically stable state.

Theorem 6 has a nice geometric interpretation, that can be seen by referring back to Figures 2.1 and 2.2 which regard Example 1. From equation 2.5, $n_a^A = 7 \leq N^A$, and therefore $(N^A, 0) \in \mathcal{V}(\omega_{aa})$. Figure 2.1 represents the basins of attraction when $k_A = k_B$. In this case, the minimum cost $\omega_{aa}$-tree has cost of 7 - the costly edge of the path being $\big(\omega_{bb} \rightarrow (7,0)\big)$. Similarly the minimum cost $\omega_{bb}$-tree has cost of 9 - the costly edge of the path being $\big(\omega_{aa}, (1,5)\big)$. Thus when $k_A = k_B$, clearly $\omega_{aa}$ is the stochastically stable outcome.

Now look at Figure 2.2 representing the basins of attraction when $\Psi = (\Psi_1^A, \Psi_2^B)$. The minimum cost $\omega_{aa}$-tree is unchanged, since the transition $\big(\omega_{bb} \rightarrow (7,0)\big)$, is still an edge on a path of minimum cost from $\omega_{bb}$ to $\mathcal{V}(\omega_{aa})$. The cost of the minimum cost $\omega_{bb}$-tree is different to when $k_A = k_B$, since now the transition $\big(\omega_{aa} \rightarrow (3,5)\big)$ is the first edge on a path of minimum cost from $\omega_{aa}$ to $\mathcal{V}(\omega_{bb})$. This has a cost of 7. So the long run distribution assigns positive probability to both $\omega_{bb}$ and $\omega_{aa}$.

Consider Figure 2.1 and suppose that $\Psi = (\Psi_1^A, \Psi_3^B)$. Note that $\omega_{ba} \in \Omega^{A,b\succ a}$ and so we use the second method described above. The example is not generic and $(N^A, N^B) = (10, 5)$, so that

$$\big((k_A, r_A), (k_B, r_B)\big) = \big((5,0), (1,2)\big)$$

Clearly then, $\hat{k} = 1$, and $\Psi^{\leftarrow \hat{k}}\big((B^0)_+\big) = \Psi^{\leftarrow \hat{k}}\big((6,0)\big) = (5,3)$, and by equation 2.20, $(B^\infty)_{NW} = (B^{\hat{n}})_{NW} = (B^2)_{NW} = (4,5)$. Describing the sets we have that

$$\begin{aligned} B^0 &= \Omega^{A,b\succeq a} \\ B^1 &= B^0 \cup \{\omega \mid [\omega]_A + [\omega]_B = 8, [\omega]_A \leq 5\}^{\downarrow} \\ B^2 &= B^1 \cup \{(4,5)\}^{\downarrow} \\ B^3 &= B^2 \end{aligned}$$

Now, we have that $\mathcal{V}(\omega_{bb}) = \Omega^{A,b\succeq a} \cup (B^{\hat{n}} \cap \Omega^{A,a\succ b}) = \{(3,5),(4,4),(5,3),(6,1)\}^{\downarrow}$.

Having classified conditions for which the various equilibria are selected, the next Section compares the selected equilibrium to those a planner might want to induce.

## 2.6 Stability versus Welfare

When the population is homogeneous, players collectively agree on what action they "would like to have taken" earlier today, and hence what they will choose for tomorrow if afforded a revision opportunity. The main issue in existing large population coordination problems is the tension between efficiency and risk-dominance. Both KMR and Young (1993) show that the risk dominant action will emerge under perturbed best response based dynamics. This result is *negative* in the sense that the locally risk-dominant equilibrium action need not coincide with the pareto dominant one, and may have payoffs that are dominated to an arbitrary extent.

In the Language Game, pareto efficiency is useless as a selection device. Both symmetric profiles are socially efficient equilibria, and there is never uniform preference over these. The purpose of this Section is to rank profiles in $E(\mathcal{G})$ according to various welfare criteria, and then compare this ranking to the outcome(s) selected by the dynamics. To infer how members of the population rank profiles in $E(\mathcal{G})$, it suffices to analyse the situation from the perspective of any one agent from each Group.

**Theorem 7.** *Within the set of group-symmetric profiles:*[23]

1. *$\omega_{aa}$ and $\omega_{bb}$ are always socially efficient.*

2. *$\omega_{ab}$ is socially efficient if and only if*

$$p \geq \frac{N-1}{2N-N^B-2} \quad \textit{and} \quad q \geq \frac{N-1}{2N-N^A-2} \tag{2.18}$$

*Proof.* The proof is straightforward and is omitted. □

A natural question to ask is the relationship between social efficiency and equilibrium. It turns out that a socially efficient profile must be an equilibrium, but not all equilibria are socially efficient.

**Theorem 8.** *If $\omega_{ab}$ is socially efficient, then it must be an equilibrium. But $\omega_{ab}$ may be an equilibrium without being socially efficient.*

*Proof.* Follows from conditions in part 4 of Theorem 3, and those in part 2 of Theorem 7. Clearly the second implies the first, while the first need not imply the second. □

Recall the examples from Section 3.3. In Example 1, with $\mathcal{G}^1 = (10, 5, 3/5, 3/5)$ and $E(\mathcal{G}^1) = \{\omega_{bb}, \omega_{aa}\}$, we have $\{U^A(\omega_{aa}), U^B(\omega_{bb})\} = \{42/5, 28/5\}$, and $\{U^A(\omega_{bb}), U^B(\omega_{bb})\} = \{28/5, 42/5\}$. The 10 Group $A$ members desire $\omega_{aa}$, while the 5 in Group $B$ prefer $\omega_{bb}$. In Example 2, with $\mathcal{G}^2 = (10, 5, 3/5, 5/6)$ and $E(\mathcal{G}^2) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, we get $\{U^A(\omega_{aa}), U^B(\omega_{aa})\} = \{42/5, 28/6\}$, $\{U^A(\omega_{ab}), U^B(\omega_{ab})\} = \{27/5, 50/6\}$, and $\{U^A(\omega_{bb}),\ U^B(\omega_{bb})\} = \{28/5, 140/6\}$. Note that $\omega_{ab}$ is pareto dominated by $\omega_{bb}$, and so provides an example of Theorem 8 at work.

The next obvious question is whether or not a decentralized adjustment process can ever select a socially-inefficient convention. In Example 2, $\omega_{ab}$ is socially inefficient but not stochastically stable. Theorem 9 shows that this is generalizable.

[23] Again, this Theorem is stated for correctly specified utility functions, and not those as given by equations (2.1)-(2.4). See footnote **??**.

**Theorem 9.** *Suppose* $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$*. For any Group-Darwinian adjustment process,* $\Psi$*, it will never be the case that* $\omega_{ab}$ *is socially inefficient and yet* $\omega_{ab} \in \Xi(\mathcal{G}, c_\Psi)$*.*

*Proof.* The proof is found in Appendix 2.9.3. □

There are a host of other commonly used social welfare criteria from the theory of social choice. Other interesting ones to apply here are the Utilitarian Welfare function, and the Rawlsian Welfare function. Under these criteria, inefficient outcomes can certainly emerge.

## 2.7 Comparative Statics

A few obvious questions spring to mind. Would a player always prefer to be part of a larger group? Would a player always prefer to be part of a group with stronger relative preference?

### 2.7.1 Varying Payoffs

The full treatment of this issue is left to future research. In this subsection I provide only an illustrative example.

**Example 6** (Increased Payoffs)**.** Let $\mathcal{G}^4 = (10, 10, 4/5, 3/5)$, and $\Psi$ be such that $\Psi^A \sim_d \Psi^B$.

Using equations (2.5)-(2.8) it can be calculated that $(n_a^A, n_b^A, n_a^B, n_b^B) = (5, 12, 16, 9)$, and so by Theorem 4, it must be that $\Xi(\mathcal{G}^4, c_\Psi) = \{\omega_{aa}\}$. Clearly, average payoffs at this point are given by $U^A(\omega_{aa}) = 4/5$ and $U^B(\omega_{bb}) = 2/5$.

Now suppose that Group $B$ payoffs are modified such that their preference for action $b$ is magnified. Precisely, suppose that $\mathcal{G}^4$ is transformed to $\hat{\mathcal{G}}^4 = (10, 10, 4/5, 4/5)$. It can now be calculated that $(n_a^A, n_b^A, n_a^B, n_b^B) = (5, 16, 16, 5)$, and so by Theorem 4, it is now the case that $\Xi(\hat{\mathcal{G}}^4, c_\Psi) = \{\omega_{ab}\}$. Average payoffs now are given by $U^A(\omega_{ab}) = U^B(\omega_{ab}) = (9/19)(4/5) < 2/5$. And so Group $B$'s payoffs have gone down.

Now suppose that the Group $B$ payoffs are modified again, with their preference for action $b$ increased further still. Precisely, suppose that $\hat{\mathcal{G}}^4$ is transformed to $\hat{\hat{\mathcal{G}}}^4 = (10, 10, 4/5, 19/20)$. Now, $(n_a^A, n_b^A, n_a^B, n_b^B) = (5, 16, 19, 2)$, and by Theorem 4, it is still the case that $\Xi(\hat{\hat{\mathcal{G}}}^4, c_\Psi) = \{\omega_{ab}\}$. Average payoffs now are given by $U^A(\omega_{ab}) = (9/19)(4/5)$ and $U^B(\omega_{ab}) = 9/20 > 2/5$. Thus Group $B$'s average payoff has increased.

### 2.7.2 Varying Group Size

## 2.8 Conclusion

This paper introduces a new coordination game in the hope of shedding some light on how behaviour might develop in societies with heterogeneous agents. The environment, the "Language Game", deviates from existing large population models in one simple but important way: there are two distinct homogeneous groups with pairwise interactions occurring both within- and across-group.

Three properties matter for equilibrium selection in the Language Game. They are (i) group size, (ii) group payoffs, and (iii) the rates at which the groups react - "group dynamics". Any agent always desires to be part of a more reactive group, but does not always long for greater numbers in their group or for more polarized preferences. While interesting, the results are not robust. That is, for a given game, assumptions on the likelihood of behavioural mutations, and the full connectedness of the population were essential.

However, the fragility of the results may not be the weakness it first appears. For example, while the results of Bergin and Lipman (1996) show that any equilibrium can be selected for appropriately defined mutations, I know of no examples where the Ising model dynamics of Blume (1993) select *different* long run equilibria to those of KMR. But they can be shown to for an open set of parameters in the Language Game (Neary, 2010b). Another major limitation of existing large population pure coordination problems, is that equilibrium selection is robust to network architecture, with uniform adoption of the locally risk-dominant action stochastically stable for any network (Peski, 2010). An immediate consequence of moving the Language Game to arbitrary networks, is that both network topology and the specifics of the dynamics matter strongly for equilibrium selection (Neary, 2010b).

Evolutionary game theory has typically focused on the many different and interesting ways in which behaviour adapts in large populations. However, in my opinion, there has been too little attention on whether or not the stage games accurately capture all situations in which large populations engage. In a companion paper, (Neary, 2010c), I define a new class of large population games called "Multiple-Group Games" (MGGs). The key feature of a MGG is that the population is partitioned into groups, with players interacting pairwise with potentially anyone from the population. The only constraint is that within-group interactions must be symmetric. Across-group interactions can be anything, as long as each player has the same strategy set available in each. This adds heterogeneity in a surprisingly tractable way. Perhaps the greatest advantage of this framework is the number of new possibilities it introduces and the number of extensions it permits.

# 2.9 APPENDIX

## 2.9.1 Lattices

Let $\omega$ and $\omega'$ be elements of $\Omega$. Write $\omega \geqslant_a \omega'$ if $[\omega]_A \geq [\omega']_A$ and $[\omega]_B \geq [\omega']_B$, and $\omega \ngeqslant_a \omega'$ if $[\omega]_A < [\omega']_A$ or $[\omega]_B < [\omega']_B$. Write $\omega >_a \omega'$ if $\omega \geqslant_a \omega'$ and $\omega \neq \omega'$, and $\omega \gg_a \omega'$ if both $[\omega]_A > [\omega']_A$ and $[\omega]_B > [\omega']_B$. States $\omega'$ and $\omega''$ are *comparable*, $\omega'' \perp_a \omega'$, if $\omega' \geqslant_a \omega''$ or $\omega'' \geqslant_a \omega'$ or both, while $\omega'$ and $\omega''$ are *incomparable*, $\omega' \,\|_a\, \omega''$, if $\omega' \ngeqslant_a \omega''$ and $\omega'' \ngeqslant_a \omega'$. The pair $(\Omega, \geqslant_a)$ is a complete lattice with *bottom* element $\omega_{bb}$ and *top* element $\omega_{aa}$.

A nonempty $\Lambda \subset \Omega$ is a *chain*, if for all $\omega', \omega'' \in \Lambda$, $\omega' \perp_a \omega''$. For a given chain $\Lambda$, define $(\Lambda)_{NE} = \{\omega \in \Lambda \,|\, \forall\, \omega' \in \Lambda,\, \omega \geqslant_a \omega'\}$, and $(\Lambda)_{SW} = \{\omega \in \Lambda \,|\, \forall\, \omega' \in \Lambda,\, \omega \leqslant_a \omega'\}$. A nonempty $\Upsilon \subset \Omega$ is an *anti-chain*, if for any $\omega', \omega'' \in \Upsilon$ with $\omega' \neq \omega''$, it is the case that $\omega' \,\|_a\, \omega''$. For a given anti-chain $\Upsilon$, define $(\Upsilon)_{SE} = \{\omega \in \Upsilon \,|\, \forall\, \omega' \in \Upsilon,\, [\omega]_A \geq [\omega']_A\}$, and $(\Upsilon)_{NW} = \{\omega \in \Upsilon \,|\, \forall\, \omega' \in \Upsilon,\, [\omega]_B \geq [\omega']_B\}$.

A *down set* is a nonempty set $D \subseteq \Omega$, where if $\omega' \in D$, and $\omega'' \in \Omega$ is such that $\omega'' \leqslant_a \omega'$, then $\omega'' \in D$. A *principal down set* is a down set of the form $\{\omega\}^{\downarrow} := \{\omega' \in \Omega \,|\, \omega' \leqslant_a \omega\}$. Similarly, an *up set* is a nonempty set $U \subseteq \Omega$, where if $\omega' \in U$ and $\omega'' \in \Omega$ is such that $\omega'' \geqslant_a \omega'$, then $\omega'' \in U$. A *principal up set* is an up set $\{\omega\}^{\uparrow} := \{\omega' \in \Omega \,|\, \omega' \geqslant_a \omega\}$. The intersection of two principal down sets is a principal down set, and $\omega' \wedge \omega''$ denotes the element, referred to as the *meet* of $\omega'$ and $\omega''$, such that $\{\omega'\}^{\downarrow} \cap \{\omega''\}^{\downarrow} = \{\omega' \wedge \omega''\}^{\downarrow}$. Similarly, $\omega' \vee \omega''$, referred to as the *join* of $\omega'$ and $\omega''$, denotes the element where $\{\omega'\}^{\uparrow} \cap \{\omega''\}^{\uparrow} = \{\omega' \vee \omega''\}^{\uparrow}$. For any set $Q \subseteq \Omega$, use $Q^{\downarrow}$ ($Q^{\uparrow}$) to denote the down (up) set generated by $Q$. That is, $Q^{\downarrow} = \cup_{\omega \in Q} \{\omega\}^{\downarrow}$.

If $\omega', \omega'' \in \Omega$ are such that $\omega' \leqslant_a \omega''$, then the *interval* $[\omega', \omega'']$ is defined as $[\omega', \omega''] := \{\omega'\}^{\uparrow} \cap \{\omega''\}^{\downarrow} = \{\omega \in \Omega \,|\, \omega' \leqslant_a \omega \leqslant_a \omega''\}$. If $\omega' \,\|\, \omega''$, then $[\omega', \omega''] = \emptyset$. A nonempty subset $Q \subseteq \Omega$ is said to be *convex* if $[\omega', \omega''] \subseteq Q$ for all $\omega', \omega'' \in Q$ with $\omega' \leqslant_a \omega''$. Both up sets and down sets are convex.

For any nonempty $Q \subseteq \Omega$, define the *lower boundary* of $Q$ by $(Q)_- := \{\omega \in Q \,|\, \{\omega\}^{\downarrow} \cap Q = \omega\}$. Similarly define the *upper boundary* by $(Q)_+ := \{\omega \in Q \,|\, \{\omega\}^{\uparrow} \cap Q = \omega\}$. The lower and upper boundaries each form an anti-chain.

For any pair $\omega', \omega'' \in Q^{\downarrow}$, say that $\omega \in Q$ is *join-irreducible* in $Q$, if $\omega' \vee \omega'' = \omega$ implies $\omega' = \omega$ or $\omega'' = \omega$. Similarly, for any pair $\omega', \omega'' \in Q^{\uparrow}$, say that $\omega \in Q$ is *meet-irreducible* in $Q$, if $\omega' \wedge \omega'' = \omega$ implies $\omega' = \omega$ or $\omega'' = \omega$. The sets of join-irreducible and meet-irreducible elements of $Q$ are denoted $(Q)_{--}$ and $(Q)_{++}$, and referred to as the *total lower boundary* and the *total upper boundary* of $Q$ respectively. $(Q)_{--}$ and $(Q)_{++}$ can each be viewed as a union of "horizontal" row chains $r_1, \ldots, r_m$ or as a union of "vertical" column chains $c_1, \ldots, c_n$, where for any $\omega_{i_k} \in c_i$ and $\omega_{j_l} \in c_j \neq c_i$, $\omega_{i_k} \,\|\, \omega_{j_l}$.

Define a metric, $\|\| : \Omega \times \Omega \to \{0, \ldots, N\}$, on $\Omega$ as follows. For any $\omega', \omega'' \in \Omega$, let $\|\omega', \omega''\| = \sum_{K=A,B} | [\omega']_K - [\omega'']_K |$. For a given set $Q \subseteq \Omega$ and state $\omega \notin Q$, interest will often focus on the state(s) in $Q$, that is (are) closest to $\omega$. That is, in the set $\left\{\omega' \in Q \,|\, \omega' \in \text{argmin}_{\hat{\omega} \in Q} \|\omega, \hat{\omega}\|\right\}$. Most importantly, for a given convex set $Q$, and an element $\omega \notin Q$ such that there exists an element $\omega' \in (Q)_{++}$ with $\omega \geqslant_a \omega'$, then $\min_{\hat{\omega} \in Q} \|\omega, \hat{\omega}\|$ is attained by some $\omega^{\star} \in (Q)_{++}$.

### 2.9.2 Constructing Basins of Attraction

Let $2^{\Omega}$ denote the power set of $\Omega$. Given a mapping $\Psi : \Omega \to \Omega$, consider the *inverse image map* $\Psi^{\leftarrow} : 2^{\Omega} \to 2^{\Omega}$ defined for every nonempty $Q \subseteq \Omega$ by $\Psi^{\leftarrow}(Q) = \{\omega \,|\, \Psi(\omega) \in Q\}$. In constructing the basins of attraction, the following result is used repeatedly.

**Lemma 6.** *The following are equivalent:*

1. *If $L \subseteq \Omega$ is a down set, then $\Psi^{\leftarrow}(L)$ is a down set*
2. *$\Psi$ is monotone.*

*Proof.* $(1 \Rightarrow 2)$ It is enough to restrict attention to principal lower sets, since if $(\{\omega_1\}^{\downarrow}, \ldots, \{\omega_n\}^{\downarrow})$ is a collection of principal lower sets, then $\cup_{j=1}^{n} \{\omega_j\}^{\downarrow}$ is a down set. For every $\omega \in \Omega$ we have that $\omega \in \Psi^{\leftarrow}\big(\{\Psi(\omega)\}^{\downarrow}\big)$. This by assumption is a lower set, so if $\omega' \in \Omega$ is such that $\omega' \leqslant_a \omega$, then $\omega' \in \Psi^{\leftarrow}\big(\{\Psi(\omega)\}^{\downarrow}\big)$. Clearly $\Psi(\omega') \leqslant_a \Psi(\omega)$ and hence $\Psi$ is monotone.

(2 ⇒ 1) Let $\omega \in \Omega$ and define $Q = \Psi^{\leftarrow}\big(\{\Psi(\omega)\}^{\downarrow}\big)$. Provided $Q \neq \emptyset$, consider some $\omega' \in Q$. Clearly for every $\omega'' \in \Omega$ with $\omega'' \leqslant_a \omega'$, we have that $\Psi(\omega'') \leqslant_a \Psi(\omega') \leqslant_a \omega$, which gives that $\omega'' \in Q$. Thus $Q$ is a down set. Finally, we note that for any $L \subseteq \Omega$, it is the case that $\Psi(L) = \Psi(\cup_{\omega' \in L}\{\omega'\}) = \cup_{\omega' \in L}\Psi(\omega')$, and combining this with the fact that the union of a collection of lower sets is a lower set yields the desired result.

□

If neither condition of Lemma 2 is satisfied, then $\Omega$ can be partitioned as $\big\{\Omega^{A,b\succ a}, \Omega^{A,a\succ b}\big\} = \big\{\Omega^{B,b\succ a}, \Omega^{B,a\succ b}\big\}$. Different monotonic Group-Darwinian adjustment processes have no affect on basins of attraction in this case. Thus both partitions are equivalent to $\{\mathcal{V}(\omega_{bb}), \mathcal{V}(\omega_{aa})\}$.

So suppose one of the sufficient conditions of Lemma 2 holds, and consider the partition of $\Omega$ $\big\{\Omega^{A,b\succeq a}, \Omega^{B,a\succeq b}, \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a})\big\}$. I show how to construct $\mathcal{V}(\omega_{bb})$ (the construction of $\mathcal{V}(\omega_{aa})$ follows analogously). There is no need to show how to construct $\mathcal{V}(\omega_{ab})$ when $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, since $\mathcal{V}(\omega_{ab}) = \Omega \backslash \big(\cup_{\omega \in \Omega_0 \backslash \{\omega_{ab}\}} \mathcal{V}(\omega)\big)$. Begin by constructing the following sequence of states:

$$B^0 := \begin{cases} \Omega^{A,b\succeq a}, & \text{genA} \\ \Omega^{A,b\succeq a} \backslash \big\{(n_a^A - 1, 0)\big\}, & \text{ngenA} \end{cases}$$

and for each $n \geq 1$, define $B^n := \Psi^{\leftarrow}\big(B^{n-1}\big)$. Finally, define $B^{\infty} := \bigcup_{n=0}^{\infty} B^n$.

**Lemma 7.** *For any monotonic Group-Darwinian dynamic* $\Psi$*,*

1. *For all* $n \geq 0$*,* $B^n \subseteq B^{n+1}$*.*

2. $\Psi^{\leftarrow}\big(B^{\infty}\big) = B^{\infty} = \mathcal{V}(\omega_{bb})$*.*

*Remark*: By part 1, the sets $\{B^n\}$ are weakly increasing in $n$. By part 2 the iterative procedure eventually it stops. That is, there exists an $\hat{n} \in \mathbb{N}$ such that for all $n \geq \hat{n}$, $B^n = B^{\hat{n}}$.

*Proof.* 1. The proof is by induction. It is true by definition for $n = 0$. By Lemma 6 and the fact that $\Psi(\omega) \leqslant_a \omega$ for each $\omega \in \Omega^{A,b\succeq a}$, it must be that for any lower set $Q \subseteq \Omega^{A,b\succeq a}$, that $Q \subseteq \Psi^{\leftarrow}(Q)$. And so the claim is true for $n = 1$. Assume it is true for $n = k$. We have

$$\begin{aligned} B^{k+2} &= \Psi^{\leftarrow}\big(B^{k+1}\big) \\ &= \Psi^{\leftarrow}\big(B^{k} \cup (B^{k+1} \backslash B^{k})\big) \\ &= \Psi^{\leftarrow}\big(B^{k}\big) \cup \Psi^{\leftarrow}\big(B^{k+1} \backslash B^{k}\big) \\ &\supseteq B^{k+1} \end{aligned}$$

where the third equality follows from the definition of $\Psi^{\leftarrow}$, and the inclusion from the inductive step. Thus the claim holds for $k + 1$.

2. The first equality is clear using part 1 and the fact that $\Omega$ is finite. The second inequality follows due to the convexity of the basins of attraction (Lemma 3).

□

The entire class of monotonic Group-Darwinian adjustment processes is too large to manage in a tractable way, and so I restrict attention to those that vary at constant rates. Any constant rate dynamic, $\Psi$, satisfies one of two properties: $\Psi^A \succeq_d \Psi^B$, or $\Psi^B \succ_d \Psi^A$. I consider each case separately.

There are three types of scenario to consider: those with $E(\mathcal{G}) = \{\omega_{bb}, \omega_{aa}\}$ and $\omega_{ab} \in \Omega^{B,a\succ b}$; those with $E(\mathcal{G}) = \{\omega_{bb}, \omega_{aa}\}$ but instead $\omega_{ab} \in \Omega^{A,b\succ a}$; and those where $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$. I ignore the second of these for reasons that will become clear soon.

1. $\Psi^A \succeq_d \Psi^B$.
   When $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, or $E(\mathcal{G}) = \{\omega_{bb}, \omega_{aa}\}$ with $\omega_{ab} \in \Omega^{B,a\succ b}$, it is clear that for any $\omega \in \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$, we have $\Psi(\omega) \in \Omega^{A,a\succ b}$. And so by Lemma 9 it must be that, $\mathcal{V}(\omega_{bb}) = D^{\infty} = B^{\infty} = B^0$

2. $\Psi^B \succ_d \Psi^A$.

Since for any $\omega \in \Big( \big\{(n_a^A - 1, 0)\big\}^{\uparrow} \setminus \big\{(n_a^A - 1, 0)\big\} \Big)$, it must be that $[\Psi(\omega)]_A \geq n_a^A$, we have

$$\mathcal{V}(\omega_{bb}) \subseteq \begin{cases} \big\{(n_a^A - 2, N^B)\big\}^{\downarrow} \cup (n_a^A - 1, 0), & \text{genA} \\ \big\{(n_a^A - 3, N^B)\big\}^{\downarrow} \cup \big\{(n_a^A - 2, 1)\big\}^{\downarrow}, & \text{ngenA} \end{cases}$$

The goal is to calculate $\big(\mathcal{V}(\omega_{bb})\big)_{++}$. To this end, I proceed in a slightly roundabout way. First it will be useful to define the following operator, $\Phi = (\Phi^A_{\xi_A}, \Phi^B_{\xi_B}) : \Omega \to \mathbb{N} \times \mathbb{N}$, where for each $K \in \{A, B\}$

$$\Phi^K(\omega) = \begin{cases} [\omega]_K - \xi_K, & \text{if } U^K(a;\omega) < U^K(b;\omega), \\ [\omega]_K\,, & \text{if } U^K(a;\omega) = U^K(b;\omega), \\ [\omega]_K + \xi_K, & \text{if } U^K(a;\omega) > U^K(b;\omega), \end{cases}$$

Clearly, $\Phi = (\Phi^A_{\xi_A}, \Phi^B_{\xi_B})$ agrees with constant rate process $\Psi = (\Psi_{\xi_A}, \Psi^B_{\xi_B})$ when the range of $\Phi$ is restricted to $\big\{0, \ldots, N^A\big\} \times \big\{0, \ldots, N^B\big\}$. Now define $D^0 = B^0$ and let us perform repeated applications of $\Phi^{\leftarrow}$ on $D^0$. That is, for all $n \leq \hat{n}$ define $D^n = \Phi^{\leftarrow}(D^{n-1}) \cap \Omega$. To help manage this, define $C^0 = (D^0)_+$, and for each $n < \hat{n}$, define the following $C^{n+1} := (D^{n+1})_+ \backslash (D^n)_+$. That $C^n$ is a nonempty anti-chain for each $n < \hat{n}$ is clear. It should also be clear that for all $n < \hat{n} - 1$, $(D^{n+1})_+ = (\cup_{j=0}^n D^j)_+$.

Now we are ready to proceed in calculating $D^{\hat{n}}$. Recall that Groups $A$ and $B$ respond at constant rates $\xi_A$ and $\xi_B$ respectively. There are two cases to consider

(a) $\omega_{ba} \in \Omega^{A, a \succ b}$.

Define $(k_A, k_B)$ as follows:

$$\big(k_A, k_B\big) = \begin{cases} \Big(\Big\lfloor \frac{n_a^A - 1}{\xi_A} \Big\rfloor, \Big\lfloor \frac{N^B}{\xi_B} \Big\rfloor\Big), & \text{genA} \\ \\ \Big(\Big\lfloor \frac{n_a^A - 2}{\xi_A} \Big\rfloor, \Big\lfloor \frac{N^B - 1}{\xi_B} \Big\rfloor\Big), & \text{ngenA} \end{cases}$$

Now define $\hat{k} = \min\{k_A, k_B\}$ and note that when $\hat{k} \geq 1$, then for all

$k \leq \hat{k}$

$$\Phi^{\leftarrow^{k}}\Big((C^0)_{SE}\Big) = \begin{cases} \Big(n_a^A - 1 - k\xi_A, k\xi_B\Big), & \text{genA} \\ \\ \Big(n_a^A - 2 - k\xi_A, k\xi_B + 1\Big), & \text{ngenA} \end{cases}$$

This element $(C^0)_{SE}$ is the key element, since for all $n \leq \hat{k}-1$, $(C^{n+1})_{SE}$ $= \Psi^{\leftarrow}\big(\{(C^n)_{SE}\}\big)$.

For all $n \leq \hat{k} = \hat{n} - 1$, $D^n = D^{n-1} \cup \Phi^{\leftarrow}(C^{n-1})^{\downarrow}$. Difficulties arise in calculating $C^{\hat{n}}$ and hence $D^{\hat{n}}$, since it must be that $\Psi^{\leftarrow}((C^{\hat{n}-1})_{SE}) = \emptyset$. Precisely, these difficulties arise at the state $\Phi^{\leftarrow}\big(\Phi^{\leftarrow^{\hat{k}}}\big((C^0)_+\big)\big)$, since by definition one of the following must be the case

$$\Big[\Phi^{\leftarrow^{\hat{k}}}\big((C^0)_+\big)\Big]_A - \xi_A < 0, \text{ or}$$
$$\Big[\Phi^{\leftarrow^{\hat{k}}}\big((C^0)_+\big)\Big]_B + \xi_B > N^B$$

So there are 3 cases to consider: $k_B < k_A$, $k_B > k_A$, and $k_B = k_A$. Before analysing each case, define the "remainders", $r_A$ and $r_B$ as follows

$$(r_A, r_B) = \begin{cases} \Big(n_a^A - 1 - \hat{k}\xi_A, N^B - \hat{k}\xi_B\Big), & \text{genA} \\ \\ \Big(n_a^A - 2 - \hat{k}\xi_A, N^B - 1 - \hat{k}\xi_B\Big), & \text{ngenA} \end{cases} \tag{2.19}$$

- $k_B < k_A$.
  In this case, $\Big[\Phi^{\leftarrow^{\hat{k}}}\big((C^0)_+\big)\Big]_A - \xi_A > 0$, but $\Big[\Phi^{\leftarrow^{\hat{k}}}\big((C^0)_+\big)\Big]_B + \xi_B$ $\geq N^B$, and so

$$\big((D^{\hat{n}})_+\big)_{NW} = \begin{cases} \Big(n_a^A - 1 - (\hat{k}+1)\xi_A, N^B\Big), & \text{genA} \\ \\ \Big(n_a^A - 2 - (\hat{k}+1)\xi_A, N^B\Big), & \text{ngenA} \end{cases}$$

- $k_B > k_A$.

In this case, $\left[\Phi^{\leftarrow\hat{k}}\left((C^0)_+\right)\right]_A - r_A \leq 0$, and $\left[\Psi^{\leftarrow\hat{k}}\left((C^0)_+\right)\right]_B + \xi_B < N^B$, and so

$$\left((D^{\hat{n}})_+\right)_{NW} = \begin{cases} \left(0, \hat{k}\xi_B + r_A\right), & \text{genA} \\ \\ \left(0, \hat{k}\xi_B + 1 + r_A\right), & \text{ngenA} \end{cases}$$

- $k_B = k_A$
  Now we have that $\left[\Phi^{\leftarrow\hat{k}}\left((C^0)_+\right)\right]_A - \xi_A < 0$, and $\left[\Phi^{\leftarrow\hat{k}}\left((C^0)_+\right)\right]_B + \xi_B > N^B$. There are two cases to consider,
  - if $r_A < r_B$, then

$$\left((D^{\hat{n}})_+\right)_{NW} = \begin{cases} \left(0, \hat{k}\xi_B + r_A\right), & \text{genA} \\ \\ \left(0, \hat{k}\xi_B + r_A + 1\right), & \text{ngenA} \end{cases}$$

  - if $r_A \geq r_B$, then

$$\left((D^{\hat{n}})_+\right)_{NW} = \begin{cases} \left(n_a^A - 1 - \hat{k}\xi_A - r_B, N^B\right), & \text{genA} \\ \\ \left(n_a^A - 2 - \hat{k}\xi_A - r_B, N^B\right), & \text{ngenA} \end{cases}$$

(b) $\omega_{ba} \in \Omega^{A,b\succeq a}$.

This time, define $(k_A, k_B)$ as follows:

$$\left(k_A, k_B\right) = \begin{cases} \left(\left\lfloor \frac{N^B}{\xi_A} \right\rfloor, \left\lfloor \frac{N^B}{\xi_B} \right\rfloor\right), & \text{genA} \\ \\ \left(\left\lfloor \frac{N^B-1}{\xi_A} \right\rfloor, \left\lfloor \frac{N^B-1}{\xi_B} \right\rfloor\right), & \text{ngenA} \end{cases}$$

Now define $\hat{k} = \min\{k_A, k_B\}$, and again note that when $\hat{k} \geq 1$, then for all $k \leq \hat{k}$

$$\Phi^{\leftarrow k}\Big((C^0)_+\Big) = \begin{cases} \Big(n_a^A - 1 - k\xi_A, k\xi_B\Big), & \text{genA} \\ \\ \Big(n_a^A - 2 - k\xi_A, k\xi_B + 1\Big), & \text{ngenA} \end{cases}$$

When $\omega_{ba} \in \Omega^{A,a \succ b}$ and $\Psi^B \succ_d \Psi^A$, there are only two cases to consider: $\hat{k} = k_B < k_A$, and $\hat{k} = k_B = k_A$. Remainders are again defined as in equation 2.19.

- $k_B < k_A$.
  Now, $\Big[\Phi^{\leftarrow \hat{k}}\big((C^0)_+\big)\Big]_A - \xi_A > n_{Aa} - 1 - N^B$, but $\Big[\Phi^{\leftarrow \hat{k}}\big((C^0)_+\big)\Big]_B + \xi_B \geq N^B$, and so

$$\big((D^{\hat{n}})_+\big)_{NW} = \begin{cases} \Big(n_a^A - 1 - (\hat{k}+1)\xi_A, N^B\Big), & \text{genA} \\ \\ \Big(n_a^A - 2 - (\hat{k}+1)\xi_A, N^B\Big), & \text{ngenA} \end{cases}$$

- $k_B = k_A$
  In this case we have that $\Big[\Phi^{\leftarrow \hat{k}}\big((C^0)_+\big)\Big]_A - \xi_A < n_{Aa} - 1 - N^B$, and $\Big[\Phi^{\leftarrow \hat{k}}\big((C^0)_+\big)\Big]_B + \xi_B > N^B$. However, it must be the case that $r_A \geq r_B$, in which case

$$\big((D^{\hat{n}})_+\big)_{NW} = \begin{cases} \Big(n_a^A - 1 - \hat{k}\xi_A - r_B, N^B\Big), & \text{genA} \\ \\ \Big(n_a^A - 2 - \hat{k}\xi_A - r_B, N^B\Big), & \text{ngenA} \end{cases} \tag{2.20}$$

It is almost immediate that $\mathcal{V}(\omega_{bb}) = B^{\hat{n}} = D^{\hat{n}} \cap \Omega^{A,a \succ b}$, and so we have computed what we set out to achieve.

### 2.9.3 Proofs

**Proof of Lemma 2.**

It is enough to show the existence of a state $\omega \in \Omega$ such that $\text{sign}\big(U^A(a;\omega) - U^A(b;\omega)\big) \neq \text{sign}\big(U^B(a;\omega) - U^B(b;\omega)\big)$. This is implied by the stronger conditions that both $n_a^B > n_a^A$ and $n_b^A > n_b^B$. I show only the first of these. Note that

$$\begin{aligned} n_a^B &= \Big\lceil q(N-2)+1 \Big\rceil \\ &\geq \Big\lceil (1-p)(N-2)+1 \Big\rceil \\ &= n_a^A \end{aligned}$$

1. Observe that the weak inequality is strict when $N$ is even.

2. When $n_a^B = \lceil \alpha \rceil$ and $n_a^A = \lceil \beta \rceil$, $\alpha, \beta \in \mathbb{R}$, a sufficient condition for $n_a^B > n_a^A$ is that $\alpha > \beta + 1$. This sufficient condition yields the required conclusion since

$$\begin{aligned} n_a^B > n_a^A \quad &\Longleftarrow \quad q(N-2)+1 > (1-p)(N-2)+1+1 \\ &\Longleftrightarrow \quad q(N-2) > (1-p)(N-2)+1 \\ &\Longleftrightarrow \quad N\big(p+q-1\big) > \big(p+q-1\big)2+1 \\ &\Longleftrightarrow \quad N > 2 + \frac{1}{p+q-1} \end{aligned}$$

**Proof of Lemma 3.**

Consider the map $\Psi^{\hat{m}} : \Omega \to \Omega$, and define an equivalence relation $\sim_\Psi$ on $\Omega$ by, $\omega' \sim_\Psi \omega'' \iff \Psi^{\hat{m}}(\omega') = \Psi^{\hat{m}}(\omega'')$. The equivalence relation $\sim_\Psi$ partitions $\Omega$ into the quotient set $\Omega/\sim_\Psi$, with $\sim_\Psi$-classes $\omega\big|_{\sim_\Psi}$. Define an order, $\leqslant_{\sim_\Psi}$, on the $\sim_\Psi$-classes in the following way: $\omega'\big|_{\sim_\Psi} \leqslant_{\sim_\Psi} \omega''\big|_{\sim_\Psi} \iff \Psi^{\hat{m}}(\omega') \leqslant_a \Psi^{\hat{m}}(\omega'')$. Now fix $\omega' <_a \omega'' <_a \omega'''$ with $\omega' \sim_\Psi \omega'''$. Then, by defining the *natural mapping*

$\jmath_{\sim_\Psi} : \Omega \to \Omega/\sim_\Psi$, it must be that

$$\jmath_{\sim_\Psi}(\omega') \leqslant_{\sim_\Psi} \jmath_{\sim_\Psi}(\omega'') \leqslant_{\sim_\Psi} \jmath_{\sim_\Psi}(\omega''') = \jmath_{\sim_\Psi}(\omega')$$

**Proof of Lemma 5.**

Cases 1 and 2 are analogous. I prove only the first.

First, note that since $\Omega^{A,b\succeq a}$ is an down-set and that $\Psi(\omega) \leqslant_a \omega$, for all $\omega \in \Omega^{A,b\succeq a}$, then the constraint that $\omega^\star \in (\Omega\backslash\Omega^{A,b\succeq a})_-$ is trivially satisfied since $(\Omega\backslash\Omega^{A,b\succeq a})$ is itself an up=set. Second, recall that $\Omega^{A,b\succ a} \subseteq \Omega^{A,b\succeq a}$, with equality generically. Even non-generically, for sufficiently large $m$, we have $\Psi^m(\omega) = \omega_{bb}$, for all $\omega \in \Omega^{A,b\succeq a}\backslash(n_a^A - 1, 0)$.

For any path $h' \in H(\omega, \Omega\backslash\Omega^{A,b\succeq a})$, define

$$g(h') := \#\Big\{\omega' \in \Omega^{A,b\succeq a} \,\Big|\, \exists\, \omega'' \in \Omega, (\omega' \to \omega'') \in h'\Big\}$$

By the uniqueness of the vertices in a path, $1 \leq g(h') \leq |\Omega^{A,b\succeq a}|$. Define

$$C(\omega;\gamma) := \min\Big\{c_\Psi(h') \,\Big|\, h' \in H(\omega, \Omega\backslash\Omega^{A,b\succeq a}),\text{ and } g(h') \leq \gamma\Big\}$$

We show that $C(\omega;\gamma)$ is attained by $h^\star$ for all $\gamma \leq |\mathcal{V}(\omega_{bb})|$.

The proof is by induction. First consider the case where $\gamma = 1$.

Clearly $\omega$ is the only state in $\Omega^{A,b\succeq a}$ which can be the initial state of an edge in $h'$, so its immediate successor, $\omega^o$ lies in $\Omega\backslash\Omega^{A,b\succeq a}$. For each $\omega \in \Omega^{A,b\succeq a}$, define the following set, $\mathcal{D}(\omega) := \big\{\omega' \,\big|\, \omega' \in \operatorname{argmin}_{\hat\omega \in (\Omega\backslash\Omega^{A,b\succeq a})_-} \|\hat\omega, \Psi(\omega)\|\big\}$. By the definition of $(\Omega\backslash\Omega^{A,b\succeq a})_-$ ,we have that $\omega^o \geqslant_a \hat\omega$, for some $\hat\omega \in (\Omega\backslash\Omega^{A,b\succeq a})_-$. So the cost of this path $h'$ is at least $c_\Psi(h') = \|\omega^o, \Psi(\omega)\| \geq \|\omega^\star, \Psi(\omega)\| = c_\Psi(h^\star)$. Clearly then,

$$c_\Psi(h^\star) = C(\omega; 1)$$

Now for the inductive step.

Assume that for some $\gamma$, $2 \leq \gamma \leq |(\Omega^{A,b\succeq a})| - 1$, and for all $\omega \in \Omega^{A,b\succeq a}$ we have that $c_\Psi(h^\star) = C(\omega;\gamma)$. Fix $\omega \in \Omega^{A,b\succeq a}$. Let $\omega^o$ be the immediate successor of $\omega$ in some path $h^o$ that is cost minimizing over all paths $h' \in H(\omega, \Omega \backslash \Omega^{A,b \succeq a})$, satisfying $g(h') \leq \gamma + 1$, i.e. that $c(h^o) = C(\omega; \gamma + 1)$. It is clear that $\omega^o \in \Omega^{A,b\succeq a} \cup \mathcal{D}(\omega)$, since otherwise

$$\begin{aligned} c_\Psi(h^o) &\geq \|\omega^o, \Psi(\omega)\| \\ &> \operatorname*{argmin}_{\hat{\omega} \in (\Omega \backslash \Omega^{A,b\succeq a})_-} \|\hat{\omega}, \Psi(\omega)\| \\ &= \|\omega^\star, \Psi(\omega)\| \\ &= c_\Psi(h^\star) \end{aligned}$$

If $\omega^o \in \operatorname{argmin}_{\hat{\omega} \in (\Omega \backslash \Omega^{A,b\succeq a})_-} \|\hat{\omega}, \Psi(\omega)\|$, then we have $c_\Psi(h^o) = c_\Psi(h^\star)$. So we will assume $\omega^o \in \Omega^{A,b\succeq a}$. By the induction hypothesis, we have that the path

$$h'' = \Big\{(\omega^o \to \omega^{\star\star})\Big\}$$

where $\omega^{\star\star} \in \operatorname{argmin}_{\hat{\omega} \in (\Omega \backslash \Omega^{A,b\succeq a})_-} \|\hat{\omega}, \Psi(\omega^o)\|$, is the path of minimum cost in the set $H(\omega^o, \Omega \backslash \Omega^{A,b\succeq a})$ in $\gamma$ or fewer steps. i.e.

$$c(h'') = C(\omega^o; \gamma)$$

Take the edge $(\omega \to \omega^o)$ and glue this on to the initial node of $h''$, $\omega^o$. Call this new path $h'''$. We have shown that $c(h''') = c(h^o)$ where

$$h''' := \{(\omega \to \omega^o)\} \cup \{\omega^o \to \omega^{\star\star}\}$$

Let us now show that $c_\Psi(t''') \geq c_\Psi(t^\star)$. There are two cases to consider:

1. $\omega \perp_a \omega^o$.

   It must be that either $\omega^o \leqslant_a \omega$ or $\omega <_a \omega^o$.

   - $\omega^o \leqslant_a \omega$.

     We have that $\Psi(\omega^o) \leqslant_a \Psi(\omega)$. Therefore, $\mathcal{D}(\omega) \subseteq \mathcal{D}(\omega^o)$, so we can choose $\omega^{\star\star} = \omega^\star$. Hence $c_\Psi(h''') = \|\Psi(\omega), \omega^o\| + \|\omega^{\star\star}, \Psi(\omega^o)\| \geq \|\omega^{\star\star}, \Psi(\omega^o)\| = \|\omega^\star, \Psi(\omega^o)\| \geq \|\omega^\star, \Psi(\omega)\| = c(h^\star)$.

- $\omega <_a \omega^o$.

  It must be that $\Psi(\omega) \leqslant_a \Psi(\omega^o) <_a \omega^o$. So, while now $\mathcal{D}(\omega^o) \subseteq \mathcal{D}(\omega)$, we can again choose $\omega^{\star\star} = \omega^\star$. In this case

$$\begin{aligned} c_\Psi(h''') &= \|\Psi(\omega), \omega^o\| + \|\omega^\star, \Psi(\omega^o)\| \\ &= \big(\omega^o_A - [\Psi(\omega)]_A\big) + \big(\omega^o_B - [\Psi(\omega)]_B\big) \\ &+ \big(\omega^\star_A - [\Psi(\omega^o)]_A\big) + \big(\omega^\star_B - [\Psi(\omega^o)]_A\big) \\ &= \big(\omega^o_A - [\Psi(\omega^o)]_A\big) + \big(\omega^o_B - [\Psi(\omega^o)]_B\big) \\ &+ \big(\omega^\star_A - [\Psi(\omega)]_A\big) + \big(\omega^\star_B - [\Psi(\omega)]_A\big) \\ &= \|\omega^o, \Psi(\omega^o)\| + \|\omega^\star, \Psi(\omega)\| \\ &\geq \|\omega^\star, \Psi(\omega)\| \\ &= c_\Psi(h^\star) \end{aligned}$$

2. $\omega \,\|_a\, \omega^o$.

   There are three cases to consider:

   - $\omega^o \leqslant_a \Psi(\omega)$.

     This implies that $\Psi(\omega^o) \leqslant_a \Psi(\omega)$, and so we have that $\mathcal{D}(\omega) \subseteq \mathcal{D}(\omega^o)$. Again, choose $\omega^{\star\star} = \omega^\star$, and so $c_\Psi(h''') = \|\Psi(\omega), \omega^o\| + \|\omega^{\star\star}, \Psi(\omega^o)\| \geq \|\omega^{\star\star}, \Psi(\omega^o)\| = \|\omega^\star, \Psi(\omega^o)\| \geq \|\omega^\star, \Psi(\omega)\| = c_\Psi(h^\star)$.

   - $\omega^o >_a \Psi(\omega)$.

     It is clear that $\operatorname{argmin}_{\hat{\omega} \in (\Omega \setminus \Omega^{A, b \succeq a})_-} \|\hat{\omega}, \omega^o\| \subseteq \mathcal{D}(\omega) \cap \mathcal{D}(\omega^o)$. So, again setting $\omega^{\star\star} = \omega^\star$, we have that

$$\begin{aligned} c_\Psi(h''') &= \|\Psi(\omega), \omega^o\| + \|\omega^\star, \Psi(\omega^o)\| \\ &= \big(\omega^o_A - [\Psi(\omega)]_A\big) + \big(\omega^o_B - [\Psi(\omega)]_B\big) \\ &+ \big(\omega^\star_A - [\Psi(\omega^o)]_A\big) + \big(\omega^\star_B - [\Psi(\omega^o)]_A\big) \\ &= \big(\omega^o_A - [\Psi(\omega^o)]_A\big) + \big(\omega^o_B - [\Psi(\omega^o)]_B\big) \\ &+ \big(\omega^\star_A - [\Psi(\omega)]_A\big) + \big(\omega^\star_B - [\Psi(\omega)]_A\big) \\ &= \|\omega^o, \Psi(\omega^o)\| + \|\omega^\star, \Psi(\omega)\| \\ &\geq \|\omega^\star, \Psi(\omega)\| \\ &= c_\Psi(h^\star) \end{aligned}$$

- $\omega^o \parallel_a \Psi(\omega)$.

  Without loss of generality we can assume that $[\Psi_\omega]_A < [\omega^o]_A$ and $[\Psi_\omega]_B > [\omega^o]_B$. If $\Psi(\omega^o)$ and $\Psi(\omega)$ cannot be ordered but $\mathcal{D}(\omega) \cap \mathcal{D}(\omega^o) \neq \emptyset$, then the result is immediate. So assume $\mathcal{D}(\omega) \cap \mathcal{D}(\omega^o) \neq \emptyset$, with $[\Psi(\omega)]_A < [\Psi(\omega^o)]_A$ and $[\Psi(\omega)]_B > [\Psi(\omega^o)]_B$. In this case $c_\Psi(\omega, \omega^o) \geq [\omega^o]_A - [\Psi(\omega)]_A > [\Psi(\omega^o)]_A - [\Psi(\omega)]_A > \big[\big(\mathcal{D}(\omega)\big)_{SE}\big]_A - [\Psi(\omega)]_A = \|\omega^\star, \Psi(\omega)\|$, where the first inequality follows by assumption, the second by monotonicity, and the third since $\mathcal{D}(\omega) \cap \mathcal{D}(\omega^o) = \emptyset$.

**Proof of Theorem 4.**

Partition the set $\Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}$ into $\{\Omega_{ab}, \bar{\Omega}_{ab}\}$, where

- $\Omega_{ab} := \{\omega \mid [\omega]_A \geq n^A_a\} \cap \{\omega \mid [\omega]_B \leq N^B - n^B_b\}$)
- $\bar{\Omega}_{ab} := \big(\Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}\big) \backslash \Omega_{ab}$.

Lemma 8 below is key in computing the minimum of $c_\Psi(h')$ and $c_\Psi(h'')$ over all $h' \in H(\omega_{ab}, \Omega^{A,b\succeq a})$ and all $h'' \in H(\omega_{ab}, \Omega^{B,a\succeq b})$. Lemma 9 shows that when both groups adopt at constant rates, the partition of $\Omega$ into regions of preference is always closely related to the basins of attraction. For reasons of simplicity, Lemma

9 is stated for the generic case only. Both proofs are omitted. That of Lemma 8 proceeds along similar lines to Lemma 5, while that of Lemma 9 is simple.

**Lemma 8.** *Suppose $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, and let $\Psi$ be a monotonic Group-Darwinian adjustment process such that both groups adapt at constant rates. Then,*

1. *For all $\omega \in \Omega_{ab}$, the minimum of $c_\Psi(h')$ over all paths $h' \in H(\omega, \Omega^{B,a\succeq b})$ is attained by*

$$h^{\star\star} := \begin{cases} \Big\{\big(\omega \to \Psi(\omega)\big), \ldots, \big(\Psi^{\hat{m}-1}(\omega) \to \Psi^{\hat{m}}(\omega)\big)\Big\} \cup \Big\{(\omega_{ab}, \omega^\star)\Big\}, \\ \{(\omega \to \omega^{\star\star})\}, \\ \Big\{\big(\omega \to \Psi(\omega)\big), \ldots, \big(\Psi^{k-2}(\omega) \to \Psi^{k-1}(\omega)\big)\Big\} \cup \Big\{(\Psi^{k-1}(\omega) \to \omega^{\star\star\star})\Big\} \end{cases}$$

$$\text{if} \begin{cases} [\Psi^m(\omega')]_A + [\Psi^m(\omega')]_B \leq N^A, \text{ for all } m \\ [\Psi(\omega)]_A + [\Psi(\omega)]_B \geq N^A \text{ and } \Psi^B \succeq_d \Psi^A \\ [\Psi^m(\omega')]_A + [\Psi^m(\omega')]_B \geq N^A, \text{ for some } m, \text{ and } \Psi^A \succ_d \Psi^B \end{cases}$$

*where,*

- $\omega^\star \in \operatorname{argmin}_{\hat{\omega}\in\Omega^{B,a\succeq b}} \|\omega_{ab}, \hat{\omega}\|$
- $\omega^{\star\star} \in \operatorname{argmin}_{\hat{\omega}\in\Omega^{B,a\succeq b}} \|\Psi(\omega), \hat{\omega}\|$
- $\omega^{\star\star\star} \in \operatorname{argmin}_{\hat{\omega}\in\Omega^{B,a\succeq b}} \|\Psi^k(\omega), \hat{\omega}\|$
- $k = \min_m \min_{\hat{\omega}\in\Omega^{B,a\succeq b}} \|\Psi^m(\omega), \hat{\omega}\|$

2. *For all $\omega \in \Omega_{ab}$, the minimum of $c_\Psi(h'')$ over all paths $h'' \in H(\omega, \Omega^{A,b\succeq a})$ is attained by*

$$h^{\star\star} := \begin{cases} \Big\{\big(\omega \to \Psi(\omega)\big), \ldots, \big(\Psi^{\hat{m}-1}(\omega) \to \Psi^{\hat{m}}(\omega)\big)\Big\} \cup \Big\{(\omega_{ab} \to \omega^\star)\Big\}, \\ \{(\omega \to \omega^{\star\star})\}, \\ \Big\{\big(\omega \to \Psi(\omega)\big), \ldots, \big(\Psi^{k-2}(\omega) \to \Psi^{k-1}(\omega)\big)\Big\} \cup \Big\{(\Psi^{k-1}(\omega) \to \omega^{\star\star\star})\Big\} \end{cases}$$

$$\text{if} \begin{cases} [\Psi^m(\omega')]_A + [\Psi^m(\omega')]_B \geq N^A, \text{ for all } m \leq \hat{m} \\ [\Psi(\omega)]_A + [\Psi(\omega)]_B \leq N^A \text{ and } \Psi^A \succeq_d \Psi^B \\ [\Psi^m(\omega')]_A + [\Psi^m(\omega')]_B \leq N^A, \text{ for some } m, \text{ and } \Psi^B \succ_d \Psi^A \end{cases}$$

*where,*

- $\omega^{\star} \in \operatorname{argmin}_{\hat{\omega}\in\Omega^{A,b\succeq a}} \|\omega_{ab}, \hat{\omega}\|$
- $\omega^{\star\star} \in \operatorname{argmin}_{\hat{\omega}\in\Omega^{A,b\succeq a}} \|\Psi(\omega), \hat{\omega}\|$
- $\omega^{\star\star\star} \in \operatorname{argmin}_{\hat{\omega}\in\Omega^{A,b\succeq a}} \|\Psi^{k}(\omega), \hat{\omega}\|.$
- $k = \min_m \min_{\hat{\omega}\in\Omega^{A,b\succeq a}} \|\Psi^{m}(\omega), \hat{\omega}\|$

**Lemma 9.** *Suppose $E(\mathcal{G}) = \{\omega_{bb}, \omega_{ab}, \omega_{aa}\}$, and let $\Psi$ be a monotonic Group-Darwinian adjustment process such that both groups adapt at constant rates. Then,*

$$\begin{aligned}
&\textit{if } \Psi^A \succ_d \Psi^B, \quad \textit{then } \mathcal{V}(\omega_{aa}) \supseteq \Omega^{B,a\succeq b}, \mathcal{V}(\omega_{bb}) = \Omega^{A,b\succeq a}, \mathcal{V}(\omega_{ab}) \subseteq \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}\\
&\textit{if } \Psi^B \succ_d \Psi^A, \quad \textit{then } \mathcal{V}(\omega_{aa}) = \Omega^{B,a\succeq b}, \mathcal{V}(\omega_{bb}) \supseteq \Omega^{A,b\succeq a}, \mathcal{V}(\omega_{ab}) \subseteq \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}\\
&\textit{if } \Psi^B \sim_d \Psi^A, \quad \textit{then } \mathcal{V}(\omega_{aa}) = \Omega^{B,a\succeq b}, \mathcal{V}(\omega_{bb}) = \Omega^{A,b\succeq a}, \mathcal{V}(\omega_{ab}) = \Omega^{A,a\succ b} \cap \Omega^{B,b\succ a}
\end{aligned}$$

We construct the minimum cost $\omega$-trees for $\omega_{bb}, \omega_{ab}$, and $\omega_{aa}$. Denote these trees of minimum cost by $\tau^{\star}_{\omega_{bb}}$, $\tau^{\star}_{\omega_{ab}}$, and $\tau^{\star}_{\omega_{aa}}$ respectively. Since the construction of $\tau^{\star}_{\omega_{bb}}$, parallels exactly the construction of $\tau^{\star}_{\omega_{aa}}$, we construct only the $\omega$-trees $\tau^{\star}_{\omega_{ab}}$ and $\tau^{\star}_{\omega_{aa}}$.

Since $\mathcal{V}(\omega_{bb})$, $\mathcal{V}(\omega_{ab})$, and $\mathcal{V}(\omega_{aa})$ are convex (Lemma 3), we have that for every pair $\omega' \in \mathcal{V}(\omega_{bb})$ and $\omega'' \in \mathcal{V}(\omega_{aa})$ such that $\omega' \perp \omega''$, there exists $\hat{\omega} \in \mathcal{V}(\omega_{ab})$ such that $\omega' <_a \hat{\omega} <_a \omega''$. Furthermore, both groups adapt at equal rates, so by lemma 9 boundaries and total-boundaries of basins of attraction coincide. We have

$$(\mathcal{V}(\omega_{bb}))_+ = (\mathcal{V}(\omega_{bb}))_{++} = \begin{cases} \left\{\omega \mid [\omega]_A + [\omega]_B = n_a^A - 1\right\}, & \text{if genA} \\ \left\{\omega \mid [\omega]_A + [\omega]_B = n_a^A - 1\right\} \backslash (n_a^A - 1, 0), & \text{if ngenA} \end{cases}$$

$$(\mathcal{V}(\omega_{ab}))_- = (\mathcal{V}(\omega_{ab}))_{--} = \left\{\omega \mid [\omega]_A + [\omega]_B = n_a^A\right\}$$

$$(\mathcal{V}(\omega_{ab}))_+ = (\mathcal{V}(\omega_{ab}))_{++} = \left\{\omega \mid [\omega]_A + [\omega]_B = N - n_b^B\right\}$$

$$(\mathcal{V}(\omega_{aa}))_- = (\mathcal{V}(\omega_{aa}))_{--} = \begin{cases} \left\{\omega \mid [\omega]_A + [\omega]_B = n_a^B\right\}, & \text{if genB} \\ \left\{\omega \mid [\omega]_A + [\omega]_B = n_a^B\right\} \backslash (n_a^B, 0), & \text{if ngenB} \end{cases}$$

- Construction of $\tau^{\star}_{\omega_{aa}}$.

For any $\omega_{aa}$-tree, $\tau_{\omega_{aa}}$, define

$$h'(\omega_{bb},\omega_{aa}) = \Big\{(\omega' \to \omega'') \in \tau_{\omega_{aa}} \,\Big|\, \omega' = \omega_{bb} \text{ or } \omega' \text{ is a successor of } \omega_{bb}\Big\}$$
$$h'(\omega_{ab},\omega_{aa}) = \Big\{(\omega' \to \omega'') \in \tau_{\omega_{aa}} \,\Big|\, \omega' = \omega_{ab} \text{ or } \omega' \text{ is a successor of } \omega_{ab}\Big\}$$

Any $\omega_{aa}$-tree, $\tau_{\omega_{aa}}$, must satisfy exactly one of the following four properties,

1. $\tau^{(1)}_{\omega_{aa}}$: $h'(\omega_{ab},\omega_{aa}) \subseteq h'(\omega_{bb},\omega_{aa})$
2. $\tau^{(2)}_{\omega_{aa}}$: $h'(\omega_{bb},\omega_{aa}) \subseteq h'(\omega_{ab},\omega_{aa})$
3. $\tau^{(3)}_{\omega_{aa}}$: $h'(\omega_{bb},\omega_{aa}) \cap h'(\omega_{ab},\omega_{aa}) = \emptyset$
4. $\tau^{(4)}_{\omega_{aa}}$: $h'(\omega_{bb},\omega_{aa}) \cap h'(\omega_{ab},\omega_{aa}) \neq \emptyset$ but neither $\tau^{(1)}_{\omega_{aa}}$ nor $\tau^{(2)}_{\omega_{aa}}$

We show that $\tau^{\star}_{\omega_{aa}}$ must possess property $\tau^{(1)}_{\omega_{aa}}$.

By Lemma 8, the minimum of $c_\Psi\big(h(\omega_{ab},\omega_{aa})\big)$ over paths in $H(\omega_{ab},\omega_{aa})$ is achieved by

$$\begin{aligned} h^\star = &\Big\{\big(\omega_{ab} \to (N^A, n_a^B - N^A)\big)\Big\} \\ &\cup \Big\{\big(\omega' \to \Psi(\omega')\big) \,|\, \omega' = \Psi^m\big((N^A, n_a^B - N^A)\big) \text{ for some } m \geq 0\Big\} \end{aligned}$$

where $c_\Psi(h^\star) = (n_a^B - N^A)$. While $h^\star$ above is different for the nongeneric case, it is easy to see that the cost $c_\Psi(h^\star)$ is unchanged.

Decompose $h'(\omega_{bb},\omega_{aa})$ into paths $h'$ and $h''$ where $h' \in H(\omega_{bb}, \Omega\backslash\Omega^{A,b\succeq a})$, and $h'' = h'(\omega_{bb},\omega_{aa})\backslash h'$, so that $c_\Psi\big(h'(\omega_{bb},\omega_{aa})\big) = c_\Psi(h') + c_\Psi(h'')$. By Lemma 5, the minimum of $c_\Psi(h')$ is attained by $h^{**} = \{(\omega_{ab} \to \omega^{\star\star})\}$, where $\omega^{\star\star} \in (\mathcal{V}(\omega_{ab}))_-$. Because $\Psi^{\hat{m}}(\omega^{\star\star}) = \omega_{ab}$ for all $\omega^{\star\star} \in \mathcal{V}(\omega_{ab})$, an upper bound for $h'(\omega_{bb},\omega_{aa})$ is given by

$$c_\Psi(h') + c_\Psi(h^\star) = n_a^A + \big(n_b^A - N^A\big) \tag{2.21}$$

Since $n_a^A \leq N^A$, and $k_A = k_B$, for all $\omega^{\star\star} \in (\mathcal{V}(\omega_{ab}))_-$ and all $m \geq 0$, it must be that $n_a^A \leq [\Psi^m(\omega^{\star\star})]_A + [\Psi^m(\omega^{\star\star})]_B \leq N^A$. If $n_a^A = N^A$, and $\Psi^k(\omega^{\star\star}) \neq \omega_{ab}$ with $\big[\Psi^k(\omega^{\star\star})\big]_A \geq N^A - n_b^B$ for some $k \geq 1$, and nongenericA, then by techniques similar to Lemma 8, the minimum of $c_\Psi\big(h(\omega_{bb},\omega_{aa})\big)$ over all $H(\omega_{bb},\omega_{aa})$ is attained by

$$\begin{aligned} h''' = &\left\{\big(\omega_{bb} \to (0,n_a^A)\big)\right\} \\ &\cup \Big\{\big(\omega' \to \Psi(\omega')\big) \,|\, \omega' = \Psi^m\big((0,n_a^A)\big) \text{ for } 0 \le m \le k-1\Big\} \\ &\cup \left\{\big(\Psi^{k-1}\big((0,n_a^A)\big) \to \hat\omega\big)\right\} \end{aligned}$$

where $\hat\omega \in \min_{\omega' \in \mathcal{V}(\omega_{aa})} \|\Psi^k\big((0,n_a^A)\big),\omega'\|$. This yields

$$c_\Psi(h''') = n_a^A + (n_b^A - N^A - 1)$$

Otherwise, it follows easily that $\min_m \min_{\hat\omega \in \Omega^{B,a\succeq b}} \|\Psi^m(\omega^{\star\star}),\hat\omega\|$ is attained at $m = \hat m$ and $\hat\omega = (N^A, N^B - n_b^B + 1)$, so that the bound in equation 2.21 binds.

- $\tau^{(1)}_{\omega_{aa}}$: This exactly attains the bound in equation 2.21.
- $\tau^{(2)}_{\omega_{aa}}$: This exactly attains the bound in equation 2.21, but only when $n_a^A = N^A$, $\Psi^k(\omega^{\star\star}) \neq \omega_{ab}$ with $\big[\Psi^k(\omega^{\star\star})\big]_A \ge N^A - n_b^B$ for some $k \ge 1$, and nongenericA all hold. Otherwise there exists $h' = h(\omega_{ab},\omega_{aa}) \setminus h(\omega_{bb},\omega_{aa})$ with $c_\Psi(h') > 0$.
- $\tau^{(3)}_{\omega_{aa}}$: In this case, it is clear that since $h(\omega_{bb},\omega_{aa})$ can be decomposed into $h'$ and $h''$ as before, with $h'' \cap h^\star \neq \emptyset$, it must be $c_\Psi\big(h(\omega_{bb},\omega_{aa})\big) > c_\Psi(h^{\star\star})$. Coupling this with the fact that the minimum cost path in $H(\omega_{ab},\omega_{aa})$ has cost equal to $c_\Psi(h^{\star\star}) = N^B - n_b^B$, yields the desired result.
- $\tau^{(4)}_{\omega_{aa}}$: Follows along similar lines to $\tau^{(3)}_{\omega_{aa}}$.

- Construction of $\tau^\star_{\omega_{ab}}$.
  For any $\omega_{ab}$-tree, $\tau_{\omega_{ab}}$, define

$$h'(\omega_{bb},\omega_{ab}) = \left\{(\omega',\omega'') \in \tau_{\omega_{ab}} \,\middle|\, \omega' = \omega_{bb} \text{ or } \omega' \text{ is a successor of } \omega_{bb}\right\}$$
$$h'(\omega_{aa},\omega_{ab}) = \left\{(\omega',\omega'') \in \tau_{\omega_{ab}} \,\middle|\, \omega' = \omega_{aa} \text{ or } \omega' \text{ is a successor of } \omega_{aa}\right\}$$

By Lemma 5 the path of minimum cost from $\omega_{bb}$ to $\Omega\backslash\Omega^{A,b\succeq a}$ is attained by $h' = (\omega_{bb}, \omega^\star)$ where $\omega^\star \in (\Omega\backslash\Omega^{A,b\succeq a})_-$. But note in this case that $(\Omega\backslash\Omega^{A,b\succeq a})_- \subseteq \mathcal{V}(\omega_{ab})$. Thus define

$$h^\star = \begin{cases} \{(\omega_{bb} \to \omega^\star)\} \\ \cup\Big\{\big(\omega', \Psi(\omega')\big) \,|\, \omega' = \Psi^m(\omega^\star) \text{ for some } m \geq 0\Big\}, & \text{if genA} \\ \\ \big\{\big(\omega_{bb} \to (n_a^A - 1, 0)\big)\big\} \cup \big\{\big((n_a^A - 1, 0) \to (n_a^A, 0)\big)\big\} \\ \cup\Big\{\big(\omega', \Psi(\omega')\big) \,|\, \omega' = \Psi^m\big((n_a^A, 0)\big) \text{ for some } m \geq 0\Big\}, & \text{if ngenA} \end{cases} \tag{2.22}$$

Both if genericA or ngenericA, $c_\Psi(h^\star) = n_a^A$. A similar analysis shows that $c_\Psi(h'')$ is minimized over all $h'' \in H(\omega_{aa}, \omega_{ab})$, by $h^{\star\star}$ where

$$h^{\star\star} = \begin{cases} \{(\omega_{aa} \to \omega^{\star\star})\} \\ \cup\Big\{\big(\omega', \Psi(\omega')\big) \,|\, \omega' = \Psi^m(\omega^\star) \text{ for some } m \geq 0\Big\}, & \text{if genB} \\ \\ \big\{\big(\omega_{aa} \to (N^A, N^B - n_b^B + 1)\big)\big\} \\ \cup\big\{\big((N^A, N^B - n_b^B + 1) \to (N^A, N^B - n_b^B)\big)\big\} \\ \cup\Big\{\big(\omega', \Psi(\omega')\big) \,|\, \omega' = \Psi^m\big((N^A, N^B - n_b^B)\big) \\ \text{ for some } m \geq 0\Big\}, & \text{if ngenB} \end{cases} \tag{2.23}$$

Again, there are 4 mutually exclusive properties that $\tau^\star_{\omega_{ab}}$ may satisfy,

1. $\tau^{(1)}_{\omega_{ab}}$: $h'(\omega_{bb}, \omega_{ab}) \cap h'(\omega_{aa}, \omega_{ab}) = \emptyset$
2. $\tau^{(2)}_{\omega_{ab}}$: $h'(\omega_{bb}, \omega_{ab}) \subseteq h'(\omega_{aa}, \omega_{ab})$
3. $\tau^{(3)}_{\omega_{ab}}$: $h'(\omega_{aa}, \omega_{ab}) \subseteq h'(\omega_{bb}, \omega_{ab})$
4. $\tau^{(4)}_{\omega_{ab}}$: $h'(\omega_{bb}, \omega_{ab}) \cap h'(\omega_{aa}, \omega_{ab}) \neq \emptyset$ but neither $\tau^{(2)}_{\omega_{aa}}$ nor $\tau^{(3)}_{\omega_{aa}}$

Recall that $\mathcal{V}(\omega_{ab})$ is *sandwiched* between $\mathcal{V}(\omega_{bb})$ and $\mathcal{V}(\omega_{aa})$, in the sense that $\mathcal{V}(\omega_{bb}) \subset \mathcal{V}(\omega_{ab})^\downarrow$ and $\mathcal{V}(\omega_{aa}) \subset \mathcal{V}(\omega_{ab})^\uparrow$. Thus, unless genA and genB and $\mathcal{V}(\omega_{ab}) = \big\{\omega \,|\, [\omega]_A + [\omega]_B = N^A\big\}$, in which case property $\tau^{(4)}_{\omega_{ab}}$ may hold due to $\mathcal{D}(\omega_{bb}) \cap \mathcal{D}(\omega_{aa}) \neq \emptyset$, then a straightforward geometric argument shows that $\tau^{\star}_{\omega_{ab}}$ must possess property $\tau^{(1)}_{\omega_{ab}}$. Regardless, combining equations 2.22 and 2.23 we get

$$c_{\Psi}(\tau^{\star}_{\omega_{ab}}) = n^A_a + n^B_b$$

Concluding,

$$\begin{aligned} c_{\Psi}(\tau^{\star}_{\omega_{bb}}) &= n^B_b + (n^A_b - N^B) \\ c_{\Psi}(\tau^{\star}_{\omega_{ab}}) &= n^B_b + n^A_a \\ c_{\Psi}(\tau^{\star}_{\omega_{aa}}) &= n^A_a + (n^B_a - N^A) \end{aligned}$$

**Proof of Theorem 5.**

Since $\omega_{ab} \in \Xi(\mathcal{G}, c_{\Psi})$ when $\Psi^A \sim_d \Psi^B$, it must be that $c_{\Psi}(\tau^{\star}_{\omega_{ab}}) \leq c_{\Psi}(\tau^{\star}_{\omega_{ab}})$ and $c_{\Psi}(\tau^{\star}_{\omega_{ab}}) \leq c_{\Psi}(\tau^{\star}_{\omega_{ab}})$. From equations (2.15)-(2.17), it must be that

$$n^A_a \leq n^A_b - N^B \tag{2.24}$$

$$n^B_b \leq n^B_a - N^A \tag{2.25}$$

and so generically

$$n^A_a \leq N^A - n^A_a + 1 \tag{2.26}$$

$$n^B_b \leq N^B - n^B_b + 1 \tag{2.27}$$

Similar to the proof of Theorem 4, I construct minimum cost $\omega$-trees of $\omega_{bb}, \omega_{ab}$, and $\omega_{aa}$ for any constant rate dynamic. Denote these by $\tau^{\star\star}_{\omega_{bb}}$, $\tau^{\star\star}_{\omega_{ab}}$, and $\tau^{\star\star}_{\omega_{aa}}$ respectively. Since the construction of $\tau^{\star\star}_{\omega_{bb}}$, parallels exactly the construction of $\tau^{\star\star}_{\omega_{aa}}$, I construct only $\tau^{\star\star}_{\omega_{ab}}$ and $\tau^{\star\star}_{\omega_{aa}}$.

- Construction of $\tau^{\star\star}_{\omega_{ab}}$.
  The result hinges on the following fact.

  *Fact* 1. The cost of $\tau^{\star}_{\omega_{ab}}$ from equation (2.16) cannot be improved upon.

  *Proof.* $\tau^{\star}_{\omega_{ab}}$ must contain paths $h^{\star}(\omega_{bb}, \omega_{ab})$ and $h^{\star}(\omega_{aa}, \omega_{ab})$.
  Consider $h^{\star}(\omega_{bb}, \omega_{ab})$. By Lemma 4 the first edge must be of the form $(\omega_{bb} \rightarrow$

$\omega^\star$) where $\omega^\star \in \Omega \backslash \Omega^{A,b\succeq a}$. Without loss of generality we can choose $\omega^\star = (n_a^A, 0)$, and the remaining edges will be a series of costless transitions to $\omega_{ab}$. Thus, $c_\Psi\big(h^\star(\omega_{bb}, \omega_{ab})\big) = n_a^A$. Similarly, the only costly edge of $h^\star(\omega_{aa}, \omega_{ab})$ an be chosen to be $(\omega_{aa} \to (N^A, N^B - n_b^B))$, so that $c_\Psi\big(h^\star(\omega_{aa}, \omega_{ab})\big) = n_b^B$.

Now $\tau^{\star\star}_{\omega_{ab}}$ must contain paths $h^{\star\star}(\omega_{bb}, \omega_{ab}) \in H(\omega_{bb}, \omega_{ab})$ and $h^{\star\star}(\omega_{aa}, \omega_{ab}) \in H(\omega_{aa}, \omega_{ab})$. Consider $h^{\star\star}(\omega_{bb}, \omega_{ab})$. It must contain a costly transition out of $\Omega^{A,b\succeq a}$. But the edge $\big(\omega_{bb} \to (n_a^A, 0)\big)$ chosen in $\tau^\star_{\omega_{ab}}$ can be chosen again. So that we can set $h^{\star\star}(\omega_{bb}, \omega_{ab}) = h^\star(\omega_{bb}, \omega_{ab})$. Similarly we can choose, $h^{\star\star}(\omega_{aa}, \omega_{ab}) = h^\star(\omega_{aa}, \omega_{ab})$.
Summarizing, we have that the minimum cost $\omega_{ab}$-tree, $\tau^{\star\star}_{\omega_{ab}}$, has cost

$$c_\Psi(\tau^{\star\star}_{\omega_{ab}}) = n_b^B + n_a^A$$

which is equal to that of $\tau^\star_{\omega_{ab}}$ as in Theorem 4. □

- Construction of $\tau^{\star\star}_{\omega_{bb}}$.
  There are two cases to consider.
  - $\Psi^A \succeq_d \Psi^B$.
    By Lemma 9, when $\Psi^A \succeq_d \Psi^B$, $\mathcal{V}(\omega_{bb}) \subseteq \Omega^{A,b\succeq a}$. Since it is now possible that $(\Omega \backslash \Omega^{B,a\succeq b})_+ \cap \mathcal{V}(\omega_{aa}) \neq \emptyset$, the minimum of $c_\Psi(\hat{h})$ over all $\hat{h} \in H(\omega_{aa}, \mathcal{V}(\omega_{bb}))$ has edges of positive cost $\big\{(\omega_{aa} \to (N^A, N^B - n_b^B))\big\}$ and $\big\{\omega_{ab} \to (n_a^A - 1, 0)\big\}$ which coincide exactly with the minimum cost $\omega_{bb}$-tree from Theorem 4, $\tau^\star_{\omega_{bb}}$, with cost given by equation 2.15.
  - $\Psi^B \succ_d \Psi^A$.
    Suppose that $\Psi = (\Psi_1^A, \mathcal{B}^B)$, so that $\Psi^B$ dominates $\Psi^A$ maximally. This maximizes $\mathcal{V}(\omega_{bb})$ so that,

$$\mathcal{V}(\omega_{bb}) = \begin{cases} \big(\big\{(n_a^A - 2, N^B)\big\}^{\downarrow} \cup \big\{(n_a^A - 1, 0)\big\}\big) \cap \Omega^{B,b\succ a}, & \text{if genA} \\ \big(\big\{(n_a^A - 3, N^B)\big\}^{\downarrow} \cup \big\{(n_a^A - 2, 1)\big\}^{\downarrow}\big) \cap \Omega^{B,b\succ a}, & \text{if ngenA} \end{cases}$$

    The cost $n_b^B + n_b^A - N^B$, attained in equation 2.15 is still attainable by the $\omega_{bb}$-tree, $\tau^\star_{\omega_{bb}}$, constructed in Theorem 4, and so $c_\Psi(\tau^\star_{\omega_{bb}})$ is an upper bound for $c_\Psi(\tau^{\star\star}_{\omega_{bb}})$. The other candidate $\omega_{bb}$-tree, $\hat{\tau}_{\omega_{bb}}$, has paths $h(\omega_{ab}, \omega_{bb})$ and $h(\omega_{aa}, \omega_{bb})$, where $h(\omega_{aa}, \omega_{bb}) \subset h(\omega_{aa}, \omega_{bb})$. The only costly transitions along $\hat{\tau}_{\omega_{bb}}$ are given by $(\omega_{ab} \to \hat{\omega})$ and $(\omega_{aa} \to \hat{\hat{\omega}})$, where $\hat{\omega} = \big((\mathcal{V}(\omega_{aa}))_-\big)_{SE}$ and $\hat{\hat{\omega}} = \big((\mathcal{V}(\omega_{bb}))_+\big)_{NW}$. Thus, $\hat{\tau}_{\omega_{bb}}$ has cost given by

$$
\begin{aligned}
c_\Psi(\hat{\tau}_{\omega_{ab}}) &= c_\Psi(\omega_{ab}, \hat{\omega}) + c_\Psi(\omega_{aa}, \hat{\hat{\omega}}) \\
&= N^B - n^B_b + 1 + \max\left\{N^A - n^A_a + 2, n^B_b\right\}
\end{aligned}
$$

There are two cases to consider:

* $n^B_b \geq N^A - n^A_a + 2$.

  Then,

$$
\begin{aligned}
c_\Psi(\hat{\tau}_{\omega_{ab}}) &= (N^B - n^B_b + 1) + n^B_b \\
&\geq (N^B - n^B_b + 1) + (N^A - n^A_a + 1) + 1 \\
&\geq n^B_b + n^A_a + 1 \\
&> c_\Psi(\tau^\star_{\omega_{ab}})
\end{aligned}
$$

  where the first inequality follows by assumption, the second using (2.26) and (2.27).

* $n^B_b < N^A - n^A_a + 2$.

  Then,

$$
\begin{aligned}
c_\Psi(\hat{\tau}_{\omega_{ab}}) &= (N^B - n^B_b + 1) + (N^A - n^A_a + 1) + 1 \\
&\geq n^B_b + n^A_a + 1 \\
&> c_\Psi(\tau^\star_{\omega_{ab}})
\end{aligned}
$$

  using (2.26) and (2.27).

**Proof of Theorem 6.**

Consider part 1 when $(N^A, 0) \in \mathcal{V}(\omega_{aa})$.

1. Construction of $\tau^\star_{\omega_{bb}}$.

   This is straightforward. The minimum cost path $h \in H(\omega_{bb}, \omega_{aa})$ must involve a transitioning out of $\Omega^{A,b\succeq a}$. Regardless of rates of evolution, it is always the case that $(n^A_a, 0) \in \mathcal{V}(\omega_{aa})$, since clearly there exists $\hat{m}$ such that for all $m \geq \hat{m}$, $\Psi^m\big((n^A_a, 0)\big) \geq (N^A, 0) \in \mathcal{V}(\omega_{aa})$. Futhermore, $(n^A_a, 0) \in \big(\Omega \backslash \Omega^{A,b\succeq a}\big)_-$, and so by Lemma 5, the path of minimum cost from $\omega_{bb}$ to $(n^A_a, 0)$ is to transition there immediately, in the generic case, and transition to $(n^A_a - 1, 0)$ and then to $(n^A_a, 0)$ in the non-generic case. Either way, this path $h(\omega_{bb}, (n^A_a, 0))$ has cost of $n^A_a$. Clearly then, a path of minimum cost from $\omega_{bb}$ to $\omega_{aa}$ is given by $h^\star := \Big\{(\omega_{bb} \to (n^A_a, 0))\Big\} \cup \Big\{\big(\omega' \to \Psi(\omega')\big) \,\big|\, \omega' = \Psi^m((n^A_a, 0)),\ \text{for some } m \geq 0\Big\}$, with $c_\Psi(h^\star) = n^A_a$.

2. Construction of $\tau^\star_{\omega_{aa}}$.

   In the same way that Lemma 5 rested on the straightforward observation that $\min_{\hat{\omega} \in \Omega \backslash \Omega^{A,b\succeq a}} \|\Psi^k(\omega), \hat{\omega}\|$ is increasing in $k$ for all $\omega \in \Omega^{A,b\succeq a}$, this subcase of Theorem 6 hinges on the following Lemma. There are two subcases to consider.

   **Lemma 10.** *Suppose $E(\mathcal{G}) = \{\omega_{bb}, \omega_{aa}\}$ and $\omega_{ab} \in \mathcal{V}(\omega_{aa})$. If $\Psi$ is a constant rate dynamic, then for all $\omega \in \mathcal{V}(\omega_{aa}) \cap \Omega^{B,b\succeq a}$, it is the case that*

   $$\min_{\hat{\omega} \in \mathcal{V}(\omega_{bb})} \|\omega, \hat{\omega}\| \;\leq\; \min_{\hat{\omega} \in \mathcal{V}(\omega_{bb})} \|\Psi(\omega), \hat{\omega}\| \tag{2.28}$$

   *Proof.* When $\Psi^A \succeq_d \Psi^B$ the result is immediate, since $\mathcal{V}(\omega_{bb}) = \Omega^{A,b\succeq a}$ by Lemma 9. So with $\Psi^B \succ_d \Psi^A$, there are two cases to consider: $\omega_{ba} \in \mathcal{V}(\omega_{bb})$ and $\omega_{ba} \notin \mathcal{V}(\omega_{bb})$.

   - $\omega_{ba} \in \mathcal{V}(\omega_{bb})$.

     Clearly $(\mathcal{V}(\omega_{bb}))_+$ can be uniquely decomposed into a collection of column chains $\{c_i\}_{i=1}^n$, where for any $\omega_{i_k} \in c_i$ and $\omega_{j_l} \in c_j \neq c_i$, it must be that $\omega_{i_k} \parallel_a \omega_{j_l}$. Now note that for any $\omega \in \mathcal{V}(\omega_{aa}) \cap \Omega^{B,b\succ a}$, it must be that $\text{argmin}_{\hat{\omega} \in \mathcal{V}(\omega_{bb})} \|\hat{\omega}, \omega\|$ is attained by either $\{\omega_{i_k}\}$ where $\omega_{i_k} \in c_i$ and $[\omega_{i_k}]_B = [\omega]_B$, or by $\{\omega_{i_k}, \omega_{j_l}\}$ where $\omega_{i_k} \in c_i$, $\omega_{j_l} \in c_j \neq c_i$, with

$[\omega_{i_k}]_B = [\omega]_B$ and $[\omega_{j_l}]_B = [\omega_{i_k}]_B - 1$ and $[\omega_{j_l}]_A = [\omega_{i_k}]_A + 1$.

For all $\omega \in \mathcal{V}(\omega_{aa}) \cap \Omega^{B,b\succ a}$ with $[\omega]_B - k_B \geq 0$, it must be that $|\,[\omega]_B - [\Psi\omega)]_B\,| > |\,[\omega]_A - [\Psi\omega)]_A\,|$ and hence (2.28) holds. If $[\omega]_B - k_B \leq 0$, the result is trivial.

We proceed to a contradiction. Suppose for the sake of simplicity that $\mathrm{argmin}_{\hat{\omega}\in\mathcal{V}(\omega_{bb})} \|\hat{\omega}, \omega\|$ is attained by $\omega_{i_k} \in c_i$ where $[\omega_{i_k}]_B = [\omega]_B$, and consider the interval $[\omega_{i_k}, \omega]$. Similarly suppose $\mathrm{argmin}_{\hat{\omega}\in\mathcal{V}(\omega_{bb})} \|\hat{\omega}, \Psi(\omega)\|$ is attained by $\omega_{j_l} \in c_i$ where $[\omega_{j_l}]_B = [\omega]_B$, and consider the interval $[\omega_{j_l}, \Psi(\omega)]$. Now note that if $\|\omega_{j_l}, \Psi(\omega)\| < \|\omega_{i_k}, \omega\|$, then $\Psi^{\leftarrow}(\omega_{j_l}) >_a \omega_{i_k}$, and hence $\omega_{i_k} \notin \mathcal{V}(\omega_{bb})$.

- $\omega_{ba} \in \mathcal{V}(\omega_{aa})$.

  The techniques used for the case where $\omega_{ba} \in \mathcal{V}(\omega_{bb})$ can again be applied to all states in $\left\{\left(N - n_a^B - 1, ((\mathcal{V}(\omega_{bb}))_{++})_{NW}\right)\right\}^{\downarrow} \cap \Omega^{A,a\succ b}$. And again a similar inductive argument to that above can be applied to those states in $\left\{\left(0, ((\mathcal{V}(\omega_{bb}))_{++})_{NW} + 1\right)\right\}^{\uparrow} \cap \Omega^{A,a\succ b}$, and so the result follows naturally.

$\square$

Armed with Lemma 10, the remainder of the proof now follows by a simple counting argument as in Theorem 5.

**Proof of Theorem 9.**

The proof uses the following two properties of the $\lceil\cdot\rceil$ function. For any $x, y \in \mathbb{R}$,

$$\lceil x\rceil + \lceil -x\rceil = \begin{cases} 0, & \text{if } x \in \mathbb{Z} \\ 1, & \text{if } x \notin \mathbb{Z} \end{cases}$$

$$\lceil x\rceil + \lceil y\rceil - 1 \leq \lceil x + y\rceil$$

It is sufficient to show that $\omega_{ab} \in \Xi(\mathcal{G}, c_\Psi)$, implies $\omega_{ab}$ is socially efficient. For $\omega_{ab} \in \Xi(\mathcal{G}, c_\Psi)$, by Theorem 4, it must be that

$$n_b^B + n_a^A = \min\Big\{n_b^B + n_b^A - N^B, n_a^A + n_a^B - N^A\Big\}$$

or equivalently, that both

$$N^A \leq n_a^B - n_b^B \tag{2.29}$$

$$N^B \leq n_b^A - n_a^A \tag{2.30}$$

Now consider equations 2.29 and 2.30 for generic parameters of $\mathcal{G}$ (we show only the case where $(1-q)N + (2q-1) \notin \mathbb{Z}$ - the other case follows along similar lines). Using equations 2.1 - 2.4, we have

$$\begin{aligned} N^A &\leq n_a^B - n_b^B \\ &= \lceil q(N-2)+1\rceil - \lceil (1-q)N + (2q-1)\rceil \\ &= \lceil q(N-2)+1\rceil + \lceil -(1-q)N - (2q-1)\rceil - 1 \\ &\leq \lceil (N-2)(2q-1)\rceil \end{aligned}$$

where the first equality followed by plugging in for $n_a^B$ and $n_b^B$, the second by the first property of $\lceil\cdot\rceil$ described above, and the final inequality by the second property of $\lceil\cdot\rceil$ along with some algebraic manipulation. Thus inequalities 2.29 and 2.30 can be restated as

$$N^A \leq \lceil (N-2)(2q-1)\rceil \tag{2.31}$$

$$N^B \leq \lceil (N-2)(2p-1)\rceil \tag{2.32}$$

Now let us check the requirements on $\omega_{ab}$ to be socially efficient. Equation 2.18 can be rearranged as

$$N^A \leq \left\lfloor \frac{1}{q}(N-1)(2q-1) \right\rfloor$$
$$\leq \left\lceil \frac{1}{q}(N-1)(2q-1) \right\rceil \tag{2.33}$$

Similarly, for $\omega_{ab}$ to be socially efficient requires

$$N^B \leq \left\lceil \frac{1}{p}(N-1)(2p-1) \right\rceil \tag{2.34}$$

Finally, note that inequalities 2.31 and 2.32 imply those in 2.33 and 2.34.

# Chapter 3

# Competing Conventions in Networked Societies

**Abstract**

This paper studies the coordination problem of a large, heterogeneous population of agents who live on a network and interact repeatedly. The heterogeneity immediately implies that varying network structure can dramatically change the set of *stochastically stable equilibria* (Foster and Young, 1990), and not just the rate that they are reached as in Ellison (1993). A simple 4-player example illustrates that the precise specification of the dynamics can also affect long run equilibrium selection, even when the population is fully connected - specifically, the uniform error dynamics of Kandori, Mailath, and Rob (1993) may select very different equilibria to the Ising model dynamics of Blume (1993).

## 3.1 Introduction

That social networks are important for economic interactions is undisputed.[1] People routinely get jobs through friends. Countries trade more regularly with neighbouring countries. And so on.

Formally, a network/graph is a set of nodes/vertices, pairs of which are connected via edges/links. In a *network game*, the vertices represent players (people, firms, countries, etc), and the edges capture the notion that certain pairs of players have payoff interdependencies. The literature on network games primarily seeks to answer the following question: given the emergence of a particular network, how does its structure impact behaviour, and in particular long run equilibrium behaviour?

One natural setting to address this question is a large population coordination problem, in which the network's edges represent 2-player games between adjacent players. The seminal paper of Kandori, Mailath, and Rob (1993) (hereafter KMR), looks at a setting where players reside on a fully connected network, and each *local-interaction* is a common $2 \times 2$ game of pure coordination.[2] KMR show that if this contest is played repeatedly, where every period players best respond to past play and occasionally make mistakes, then uniform adoption of the locally risk-dominant action is the most likely, so-called *stochastically stable* (Foster and Young, 1990), equilibrium.[3,4] Ellison (1993) extends this result to some particular network structures, e.g. the circle. Using the more general framework of Morris (2000), Peski (2010) extends this *equilibrium selection* result to arbitrary networks, showing that uniform adoption of the locally risk-dominant action is always a stochastically stable equilibrium, and uniquely so if the network satisfies

[1] The literature on networks, even in economics, is far too vast to mention fully here. See textbooks by Goyal (2007) and Jackson (2008)

[2] Actually the model of KMR is one of random matching, but such a framework is strategically equivalent to that where agents live on a fully connected network and interact with everyone else.

[3] Young (1993) introduces general techniques that allow computation of stochastically stable equilibria for general games.

[4] Caveat: KMR assume that players err with state independent probabilities. While risk-dominance is also the prediction when players make logit-type errors (Blume, 1993), Bergin and Lipman (1996) showed that such predictions are highly dependent on how players make mistakes, and that any strict equilibrium can be stochastically stable for an appropriately defined model of mistakes.

a mild density condition.[5]

This seems to suggest that if one accepts stochastic stability with "intuitive" errors as the correct selection criterion for large population coordination problems, then it is not possible to implement a particular outcome by manipulating network architecture. In this paper I show that this is false. Network architecture can affect equilibrium selection without resorting to "unintuitive" mistakes, and in fact, the risk-dominance selection results of the papers mentioned above are primarily driven by another primitive of the game: the homogeneity of the population. A homogeneous population means that every edge in the network represents a common local-interaction, which by definition is constrained symmetric. This is limiting, since when local-interactions are pure coordination games, each player's coordination problem is sufficiently similar, and hence, in the presence of neighbours who are persistently making mistakes, so is optimal behaviour (i.e. take the risk-dominant action).

The heterogeneous framework I use is the "Language Game" of Neary (2010a). The Language Game deviates from existing large population models in one simple but important way: the population is partitioned into two homogeneous groups with local-interactions occurring both within- and across-group.[6] Each local-interaction is still a $2 \times 2$ game of coordination, though players from different groups prefer to coordinate on different actions. When arranged on an arbitrary network, players from the same group are differentiated only by the identity of those to whom they are connected.

Such heterogeneity introduces tension. In the homogeneous agent case, it is clear what all players would like to coordinate on - the unique pareto efficient equilibrium - the issue of interest is whether or not decentralized play will allow such coordination to occur. In the Language Game, players from different groups have different "ideal" equilibria.

I examine the stochastically stable equilibria to the Language Game on

[5]This holds for a class of noisy dynamics that includes that of KMR and Blume (1993) - though clearly cannot be the case for all such noisy dynamics (see the caveat in footnote 4).

[6]The Language Game is one of many Multiple-Group Games (MGGs) introduced in Neary (2010c).

arbitrary networks using the logit-error model of Blume (1993). To do this I exploit some existing results. The Language Game is a potential game (Shapley and Monderer, 1996).[7] This is useful, since the equilibria that maximize the potential function are stochastically stable under these noisy dynamics (see Young (2001) and Neary (2010c)). Armed with this, it can be shown that computation of the stochastically stable equilibria reduces to solving a linear programming problem. The key factor for equilibrium selection is the interplay between strength of payoffs and network structure. The stronger a particular group's preferences, the more difficult it is to dislodge them from taking their more preferred action (since mistakes are payoff dependent). The more agents from their own group to whom an agent is connected, the less likely it is that they will adopt the other group's preferred action.

The results allow something to be said about the strength of networks. In pure graph theory, there are many notions of the *centrality* of a vertex, where centrality is often interpreted as importance,[8] but each of these notions is static and simply determined by the underlying network architecture. Recently, the game theory literature has begun to examine which vertices are most important when interactions along links are strategic.[9] In each case, the focus is on which players are "stronger", with "strength" indices captured by a player's equilibrium payoff. While the Language Game is also a little abstract, it does allow me to address questions like: is it possible for a smaller, "better"-connected group to impose their preferred equilibrium outcome on a larger, "worse"-connected group? And if so, what properties of the network allow this? If a group could strategically

[7]Potential games have a simple and intuitive interpretation. Whenever a player unilaterally deviates from the current profile, the change in that person's utility is equal to the change in potential.

[8]Some of these are: *degree centrality* - the fraction of other vertices a particular vertex is connected to; *betweenness centrality* - loosely defined as the number of shortest paths between all pairs of other vertices that the vertex of interest lies on; *closeness centrality* - defined as the arithmetic mean of the closest path between the vertex and all other vertices; *eigenvector centrality* - where, loosely put, the importance of a vertex is related to the importance of its neighbouring vertices. See Jackson (2008) (pages 37-43) for reference.

[9]Polanski (2007), and Manea (2009) analyse similar models, in which a link between two players represents a bargaining opportunity over a unit surplus. Both **?** and Bramoulle and Kranton (2007) look at games where a homogeneous population of players interact pairwise with each of their neighbours via a common symmetric anti-coordination game.

arrange itself, how would it best do so in order to optimize the likelihood that its preferred outcome will be adopted?

Section 4.2 shows that network structure can dramatically affect stochastically stable outcomes, even in a simple 4-player game. It further emphasizes that the details of the dynamics are far more important for equilibrium selection in games with heterogeneity than in games with a homogeneous population, as the dynamics of KMR and Blume (1993) can select very different equilibria.

Section 3.3 formally defines the Language Game, while Section 3.4 introduces the dynamics and relates stochastic stability to the potential function.

Section 3.5 defines the linear programming problem that computes the stochastically stable outcome for any network.

While the linear programming method provides a complete solution for any network, to help fix ideas it is useful to examine the selection results for some oft-studied cases. Section 3.6 looks at equilibrium selection when the network is fully connected. In Neary (2010a), I performed a similar analysis using the uniform error dynamics of KMR, and I showed that certain group-properties, *group size*, *group payoffs*, and *group adaptiveness*, affect stochastically stable outcomes.[10] The first two properties remain important for equilibrium selection, though the property of group adaptiveness is dropped as it is not definable under the dynamics of Blume (1993). Strength of payoffs have a greater effect in this case, but this is only natural since the dynamics are payoff dependent.

In Section 3.7, I discuss stochastic stability for some other commonly studied graph structures that have received much attention in the literature. In particular, using situations where payoffs are *mirrored*, it emphasizes how in using the potential approach, it is the number of edges within each group that matter.

Section 3.8 characterizes properties of the graph such that each group will adopt their preferred equilibrium action, and thus fail to coordinate with the other group. The key feature is the way in which sets of agents, and not just individual agents, are connected to those in the other group.

Section 3.9 concludes.

[10]This was contrasted with the model of KMR, as these properties have no bearing on selection in their homogeneous agent framework.

## 3.2 Example: Network structure and dynamics matter

The purpose of this Section is to show, via a simple 4-player example, that (i) existing equilibrium selection results for network games are driven by the fact that the pairwise interaction occurring across each link is the same, and (ii) the dependence of selection results on the details of the dynamics is far more pronounced in a heterogeneous population setting.

Consider the following story. There is a population of 4 players, $\mathcal{N} = \{1, 2, 3, 4\}$, partitioned into two groups of identical agents, $\Pi = \{A, B\} = \{\{1, 2\}, \{3, 4\}\}$. Each player is located at a vertex of a fully connected undirected graph, $\Gamma$, with vertex set $\mathcal{N}$, and edge set $E$.

The strategic situation is a simultaneous move game where each player has the same two-action strategy set $\{a, b\}$. Along edge $(i, j)$, $i < j$, players $i$ and $j$ play game $G^{\pi(i)\pi(j)}$, where $\pi(k)$ is the group to which player $k$ belongs. Each player must use the same action in all local-interactions and his/her utility is the unweighted sum of payoffs earned with each neighbour. Figure 3.1 below shows $\Gamma$ and the local-interactions occurring across each edge, while payoffs are as given in Figure 3.2. The vertices for players in Group $A$ are coloured red, and for those in Group $B$ are coloured blue.

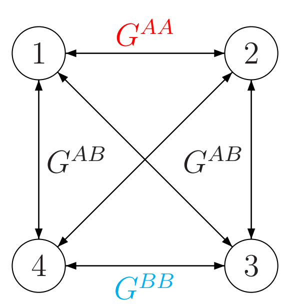


**Figure 3.1**: Fully connected graph, $\Gamma$.

Interest lies in pure strategy equilibria since these are the only serious candidates for long run behaviour. Letting $\mathbf{s} = (s_1, s_2, s_3, s_4) \in \mathbf{S} := \{a, b\}^4$ denote a particular strategy profile, the only two pure strategy equilibria to this game are

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $3, 3$ | $0, 0$ |
| $b$ | $0, 0$ | $2, 2$ |

$G^{AB}$

| $A$ \ $B$ | $a$ | $b$ |
|---|---|---|
| $a$ | $3, 2$ | $0, 0$ |
| $b$ | $0, 0$ | $2, 3$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $2, 2$ | $0, 0$ |
| $b$ | $0, 0$ | $3, 3$ |

**Figure 3.2**: $G^{AA}, G^{AB}, G^{BB}$.

$(a, a, a, a)$ and $(b, b, b, b)$, which I denote **a** and **b** respectively.

The above game gets repeated. Players follow a simple updating rule. Specifically, whenever a player is afforded a revision opportunity, he/she takes a best response against the current population profile.

Aggregating these individual decisions generates population dynamics. It is clear that the initial profile, or *state of play*, has great bearing on the terminal resting point. The pioneering works of Foster and Young (1990), Young (1993), and Kandori, Mailath, and Rob (1993) showed that by adding "noise" to the dynamics, it is possible to remove this path dependence.[11] There are two benchmark models of noise, which is introduced at the level of the individual by supposing that players occasionally make mistakes. The first is the uniform error model of KMR, where every player mistakenly chooses an inferior action with state-independent probability. The second is the logit error model of Blume (1993), whereby the likelihood of an error is proportional to the loss in utility due to the error.

Let us examine these two different scenarios. First, let us say that every period all players best respond to the current profile and make uniform errors. Second, imagine that each period one player is randomly selected and chooses his/her action according to the logit-error model. It turns out that since both groups are the same size, the graph is fully connected, and payoffs (in a sense made precise in Neary (2010c)) are *mirrored*, that for these two noisy dynamics, stochastic stability predicts both **a** and **b**, each with probability 0.5. This is hardly

[11] Adding noise means that population behaviour is always moving. However, despite the fact that population behaviour is always in flux, it spends the bulk of time centralized on a subset of the equilibria known as the "stochastically stable" equilibria.

surprising.

So far so good. Now suppose that the original fully-connected graph, $\Gamma$, is altered by removing edges $(1,3)$ and $(1,4)$. We are then left with the subgraph of $\Gamma$, $\hat{\Gamma}$, with the same vertex set $\mathcal{N}$, and set of edges $\Big\{(1,2),(2,3),(2,4),(3,4)\Big\}$. This new graph $\hat{\Gamma}$ is depicted in Figure 3.2 below. (Whatever edges remain have the same local-interaction games played across them.)

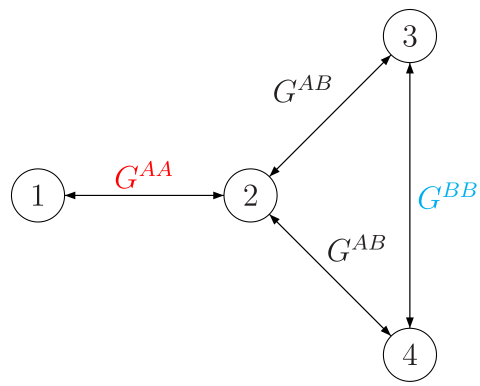


**Figure 3.3**: Network $\hat{\Gamma}$, a subnetwork of $\Gamma$.

The set of pure strategy equilibria to this new game is the same as in the fully-connected case: **a** and **b** (this can be seen by noting that player 1 must coordinate with player 2). So despite the fact that the network structure is now quite different and has lost it air of symmetry, the symmetric profiles are the only pure strategy equilibria and hence the only candidate rest points of a best-response based dynamic.

Now let us compute the stochastically stable equilibria for each of the noisy dynamics described above. It turns out that the selection results are very different.

### Best-Reply Dynamic with uniform errors

To calculate the long run distribution, I will first look at each player's updating rule for the next period. I can then, via brute force, calculate how the best-reply dynamic proceeds from each state. With this information it is then easy, but tedious, to perform a crude, but standard, "mutation counting" argument to conclude which outcome(s) is (are) stochastically stable.

The myopic updating rules for each player are as follows:

**player 1:** take whatever action player 2 took this period.

**player 2:** take whatever action was used by at least two of your neighbours this period.

**players $3, 4$:** play $b$ if at least one neighbour played $b$ this period. Otherwise play $a$.

There are $2^4 = 16$ states. Table 3.1 below lists each state, ordered lexicographically under the relation $a < b$, and examines where the best-reply dynamic terminates for each initial state, tracing the path by which it arrives there using the updating rules given above. Let $\mathbf{s}^0$ denote the initial state and $\mathcal{B}$ the best-reply dynamic. $\mathcal{B}^2(\mathbf{s}) = \mathcal{B}\big(\mathcal{B}(\mathbf{s})\big)$ is the 2-fold iteration of $\mathcal{B}$, and so on. The successor state in a sequence follows a "$\rightarrow$" symbol.

**Table 3.1**: Evolution of $\mathcal{B}$ for each initial state.

| | $\mathbf{s}^0$ | | $\mathcal{B}(\mathbf{s}^0)$ | | $\mathcal{B}^2(\mathbf{s}^0)$ | | $\mathcal{B}^3(\mathbf{s}^0)$ | | end |
|---|---|---|---|---|---|---|---|---|---|
| 1 | $(a,a,a,a)$ | $\rightarrow$ | $\mathbf{a}$ | | | | | | $\mathbf{a}$ |
| 2 | $(a,a,a,b)$ | $\rightarrow$ | $(a,a,b,a)$ | $\rightarrow$ | $(a,a,a,b)$ | $\rightarrow$ | $\cdots$ | | cycle |
| 3 | $(a,a,b,a)$ | $\rightarrow$ | $(a,a,a,b)$ | $\rightarrow$ | $(a,a,b,a)$ | $\rightarrow$ | $\cdots$ | | cycle |
| 4 | $(a,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | | | $\mathbf{b}$ |
| 5 | $(a,b,a,a)$ | $\rightarrow$ | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | $\mathbf{b}$ |
| 6 | $(a,b,a,b)$ | $\rightarrow$ | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | $\mathbf{b}$ |
| 7 | $(a,b,b,a)$ | $\rightarrow$ | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | $\mathbf{b}$ |
| 8 | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | | | | | $\mathbf{b}$ |
| 9 | $(b,a,a,a)$ | $\rightarrow$ | $\mathbf{a}$ | | | | | | $\mathbf{a}$ |
| 10 | $(b,a,a,b)$ | $\rightarrow$ | $(a,b,b,a)$ | $\rightarrow$ | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ |
| 11 | $(b,a,b,a)$ | $\rightarrow$ | $(a,b,a,b)$ | $\rightarrow$ | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ |
| 12 | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | | | $\mathbf{b}$ |
| 13 | $(b,b,a,a)$ | $\rightarrow$ | $(b,a,b,b)$ | $\rightarrow$ | $(a,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | $\mathbf{b}$ |
| 14 | $(b,b,a,b)$ | $\rightarrow$ | $\mathbf{b}$ | | | | | | $\mathbf{b}$ |
| 15 | $(b,b,b,a)$ | $\rightarrow$ | $\mathbf{b}$ | | | | | | $\mathbf{b}$ |
| 16 | $(b,b,b,b)$ | $\rightarrow$ | $\mathbf{b}$ | | | | | | $\mathbf{b}$ |

Of the 16 strategy profiles, the only rest points of $\mathcal{B}$ are the symmetric ones. States '2' and '3' form their own closed cycle. States '1' and '9' lead to $\mathbf{a}$, while the remaining 12 states all come to rest at $\mathbf{b}$.

To discuss how the stochastically stable equilibria are computed for this situation it is necessary to introduce a little bit of terminology.[12] The deterministic best-reply dynamic, $\mathcal{B}$, induces a nonergodic Markov process, with transition matrix $P$, on the state space $\mathbf{S}$. The process $P$ has three recurrent classes: $R_1 = \{\mathbf{a}\}$, $R_2 = \{\mathbf{b}\}$, and the closed cycle $R_3 = \{(a,a,a,b),(a,a,b,a)\}$. All other states will transition with probability 1 to either $\mathbf{a}$ or $\mathbf{b}$, with the set that lead to $\mathbf{a}$ ($\mathbf{b}$) referred to as the *basin of attraction* of $\mathbf{a}$ ($\mathbf{b}$).

An $\mathbf{s}$-tree on $\mathbf{S}$ is a collection of arrows between pairs of elements of $\mathbf{S}$, such that every element except $\mathbf{s}$ is the initial point of exactly one arrow, and from any state in $\mathbf{S}\setminus\{\mathbf{s}\}$, there is a unique path leading to $\mathbf{s}$.

The *cost* between any two states $\mathbf{s}_1$ and $\mathbf{s}_2$, denoted $c(\mathbf{s}_1, \mathbf{s}_2)$, is the number of simultaneous mutations needed to transform $\mathcal{B}(\mathbf{s}_1)$ to $\mathbf{s}_2$. The path of minimum cost between two states is the natural extension of this. The stochastically stable states are those contained in the recurrent class with minimum cost $\mathbf{s}$-tree.

In this example, for each recurrent class, the $\mathbf{s}$-tree of minimum cost is never unique. To compute the minimum cost $\mathbf{s}$-trees, all we need to consider are paths of minimum cost between the recurrent classes.[13] And to do this, we just have to examine paths from each recurrent class to the basin of attraction of the other recurrent classes. Table 3.2 below displays a minimum cost path and its costs for all pairs of recurrent classes.

**Table 3.2**: Transition Costs

| Transition | Cost | Via the path |
|---|---|---|
| $R_1 \to R_2$ | 1 | $(a,a,a,a) \to (a,a,a,b)$ |
| $R_1 \to R_3$ | 1 | $(a,a,a,a) \to (a,b,a,a)$ |
| $R_2 \to R_1$ | 1 | $(a,a,a,b) \to (a,a,a,a)$ |
| $R_2 \to R_3$ | 1 | $(a,a,a,b) \to (a,a,b,b)$ |
| $R_3 \to R_1$ | 3 | $(b,b,b,b) \to (b,a,a,a)$ |
| $R_3 \to R_2$ | 3 | $(b,b,b,b) \to (a,a,a,b)$ |

Write $c\big((R_i)\big)$ for the cost of a minimum cost $\mathbf{s}$-tree for recurrent class $R_i$,

[12] These techniques were first introduced to game theory in Foster and Young (1990) and are now very standard.

[13] The minimum cost $\mathbf{s}$-tree for all states in a given recurrent class is equal since transitions within a recurrent class are by definition costless - see Young (1993).

$i \in \{1, 2, 3\}$. It can be computed using Table 3.2 that

$$c\big((R_1)\big) = 4$$
$$c\big((R_2)\big) = 4$$
$$c\big((R_3)\big) = 2$$

Note then that the stochastically stable equilibrium is $(\mathbf{b}, \mathbf{b})$, which is very different to the $50 : 50$ distribution over $(\mathbf{a}, \mathbf{a})$ and $(\mathbf{b}, \mathbf{b})$, as it was when the network was fully connected. Network structure has thus affected the stochastically stable equilibrium.

**Best response dynamic with logit errors**

Now, each period, one player is drawn randomly from the population and afforded the opportunity to update his/her current action. Noise is introduced by supposing that the likelihood of playing a particular action is exponentially related to its expected payoff (and hence not independent of the current state).

Under such a dynamic, the only feasible one-period transitions are to states that differ in a maximum of one component. Drawing out a table such as Table 3.1 is not helpful since each state has multiple possible successors and costly transitions are not of equal cost. Fortunately, a convenient shortcut is available. Since each local-interaction of the Language Game is a potential game (Shapley and Monderer, 1996), the Language Game is itself a potential game (Section 3.4). It turns out that, under the noisy dynamic described above, the stochastically stable states are precisely those that maximize the potential (Theorem 10).

In this case, the potential function takes the same value at either symmetric equilibria. The reason for this is that the potential function ignores basins of attraction as it values only the potential along each edge of the network. There are the same number of within-group interactions for either group, and since payoffs are mirror image the sum of potentials for all within-group interactions is equal regardless of which symmetric profile is adopted. Furthermore, for the same reason, the potential along edges across-group takes the same value for either type of successful coordination. As such, again the stochastically stable distribution places probability of 0.5 on each symmetric equilibrium.

### Why the difference in selection for different dynamics?

Given the heterogeneity of the framework, I hope it is unsurprising that network architecture can affect equilibrium selection. It is however, at least to me, far more of a mystery as to why the different dynamics can select such radically different equilibrium outcomes. In the example outlined above, the reason seems to be that the discontinuities in selection occur at different points.

It is perhaps useful for a moment to recall the homogeneous agent setting, with a population that is fully connected. In that case, once one action risk-dominates another, it is immediate that uniform adoption of this action will be selected. Such a statement is true for either the uniform error or the logit error dynamic.

Now consider the game above on the fully connected graph. Under the uniform error dynamic, we can perturb the payoffs of the players and the prediction will not change. This is true no matter who's payoff we perturb and no matter how we perturb them, and is true because even with nearby payoffs, it is still the case that each player needs to see at least 2 others taking his preferred action in order to adopt it. But if we perturb only the payoffs of one group, say Group $A$, such that their payoffs from coordinating on their preferred equilibrium outcome are increased from 3 to $3+\varepsilon$, ($\varepsilon > 0$), then the prediction changes. The reasoning is that with payoff dependent logit errors, it is now slightly harder, albeit only fractionally so, that Group $A$ players will accidentally choose action $b$ when $a$ is their best response, than it is for Group $B$ players to accidentally choose action $a$ when $b$ is their best response.

## 3.3 The Language Game

A *network* is an undirected graph $\Gamma = (\mathcal{N}, E)$ where $\mathcal{N} = \{1, \ldots, N\}$ denotes the set of vertices, and $E \subseteq \{(i,j) | i \neq j \in \mathcal{N}\}$ denotes the set of edges, such that $(j,i) \in E$ whenever $(i,j) \in E$.[14] In the model, $\mathcal{N}$ will also denote the player set, so that each vertex is the home of a distinct player. If $(i,j) \in E$, then edge $(i,j)$ has *ends* $i$ and $j$, and vertices $i$ and $j$ are said to be *adjacent*. I abuse notation by writing $E = \{g_{ij}\}_{i,j\in\mathcal{N}}$, where $g_{ij} = 1$ if $i$ and $j$ are adjacent, and 0 otherwise.

A *path* in a graph is a sequence of distinct vertices such that each vertex is adjacent to the next vertex in the sequence. If it is possible to establish a path from any vertex to any other vertex, then the graph is said to be *path-connected*; otherwise, the graph is *disconnected*.

An undirected graph $\Gamma' = (\mathcal{N}', E')$ is said to be a *subnetwork* of network $\Gamma = (\mathcal{N}, E)$, if $\mathcal{N}' \subseteq \mathcal{N}$ and $E' \subseteq E$. For any network $\Gamma = (\mathcal{N}, E)$, and for any two nonempty subsets, $X, Y \subseteq \mathcal{N}$, we may consider the set of links with one end in $X$ and the other in $Y$, $E^{XY} := \{(i,j) \in E | i \in X, j \in Y\}$. We may thus consider the subnetwork of $\Gamma$ induced by a vertex set $C \subseteq \mathcal{N}$, and denote it as $\Gamma[C] = \left(C, E^{CC}\right)$. Similarly we may define a subgraph induced by a set of edges $F \subseteq E$, that has edge set $F$ and vertex set consisting of all vertices which lie at the end of at least one edge in $F$. This we denote by $\Gamma[F]$.

The Language Game is defined as the tuple $\mathcal{L} := \left\{\mathcal{N}, \Pi, S, \Gamma, \mathbb{G}\right\}$, where $\Pi := \{A, B\} = \left\{\left\{1, \ldots, N^A\right\}, \left\{N^A + 1, \ldots, N^A + N^B\right\}\right\}$ is a partition of $\mathcal{N}$ into nonempty *groups* $A$ and $B$ of sizes $N^A$ and $N^B$ respectively ($N = N^A + N^B$), $S := \{a, b\}$ is the strategy set common to all players, $\Gamma := (\mathcal{N}, E)$ is the network on which the players live with $\Gamma[A]$ and $\Gamma[B]$ both path-connected,[15] and $\mathbb{G} := \left\{G^{AA}, G^{AB}, G^{BB}\right\}$ is the collection of pairwise *local-interactions*, where $G^{AA}$ is the game that a player from Group $A$ plays with a player from Group $A$, etc. In $G^{AA}$ ($G^{BB}$), the two group $A$ ($B$) players are denoted $A_1$ and $A_2$ ($B_1$ and $B_2$). These games are symmetric. The game $G^{AB}$ is not symmetric and so it is important to note who is "row" and who is "column": $A_i$ denotes the row player from Group $A$, while $B_j$ denotes the column players from Group $B$. Payoffs are given as follows,

[14] The symbol $\subseteq$ ($\subset$) connotes weak (strict) inclusion,

[15] There is no steadfast reason to insist that both $\Gamma[A]$ and $\Gamma[B]$ are path-connected. Rather, if they were not path-connected, then in some sense each path-connected component of these subgraphs could be thought of as comprising separate groups.

$G^{AA}$

| $A_1$ \ $A_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $\gamma_A, \gamma_A$ | $0, 0$ |
| $b$ | $0, 0$ | $1-\gamma_A, 1-\gamma_A$ |

$G^{BB}$

| $B_1$ \ $B_2$ | $a$ | $b$ |
|---|---|---|
| $a$ | $1-\gamma_B, 1-\gamma_B$ | $0, 0$ |
| $b$ | $0, 0$ | $\gamma_B, \gamma_B$ |

$G^{AB}$

| $A_i$ \ $B_j$ | $a$ | $b$ |
|---|---|---|
| $a$ | $\gamma_A, 1-\gamma_B$ | $0, 0$ |
| $b$ | $0, 0$ | $\gamma_B, 1-\gamma_A$ |

Define $\mathbf{S} := \prod_{j=1}^{N} S_j$ with typical element $\mathbf{s} = (s_1, \ldots, s_N)$. A pure strategy profile, $\mathbf{s} \in \mathbf{S}$, can be viewed as $(s_i; \mathbf{s})$. Thus, $(\hat{s}_i; \mathbf{s}) = (s_1, \ldots, \hat{s}_i, \ldots, s_N)$. Writing $1_{\{\}}$ for the indicator function for event $\{\}$, for typical players $i \in A$ and $j \in B$, the utility from taking action $s \in S$ is given by,

$$U_i^A(s; \mathbf{s}) := \gamma_A 1_{\{s=a\}} \sum_{j \neq i} g_{ij} 1_{\{s_j=a\}} + (1-\gamma_A) 1_{\{s=b\}} \sum_{j \neq i} g_{ij} 1_{\{s_j=b\}} \tag{3.1}$$

$$U_k^B(s; \mathbf{s}) := (1-\gamma_B) 1_{\{s=b\}} \sum_{l \neq k} g_{kl} 1_{\{s_l=b\}} + \gamma_B 1_{\{s=a\}} \sum_{l \neq k} g_{kl} 1_{\{s_l=a\}} \tag{3.2}$$

The Language Game is a simultaneous move game, in which players do not randomize. Utilities are the unweighted sum of payoffs earned from interacting with neighbours, where the same action must be used with each. Note that all local-interactions are *opponent independent*, in that a player's payoff depends only on the actions chosen and not the neighbour's identity. Thus a player cares only about the number of his/her neighbours who choose the same action, and not on who those others are. I assume that $\gamma_A, \gamma_B \in (1/2, 1)$, so that regardless of who they interact with, Group $A$ members prefer to coordinate on $a$, while those in Group $B$ prefer to coordinate on $b$.

A graph, $\Gamma$, is said to be bipartite if the vertices can be divided into two disjoint sets $\mathcal{N}'$ and $\mathcal{N}''$ such that no two vertices from the same set are adjacent. The pair $(\mathcal{N}', \mathcal{N}'')$ is known as the bipartition of $\Gamma$. In the Language Game, it is clear that for all $X \subseteq A$ and $Y \subseteq B$, $\Gamma[E^{XY}]$ is bipartite. This bipartition need not be unique unless $\Gamma[E^{XY}]$ is path-connected. However, $\Gamma[E^{AB}]$ can always be taken as the bipartite graph $(X, Y, E^{AB})$ for some $X \subseteq A$ and $Y \subseteq B$, and this is the bipartition I will always adopt.

## 3.4 Dynamics and Potential

When they exist, strict equilibria are typically the most serious candidates for long run behaviour when agents are myopic best responders. In the Language Game, there are always at least two strict equilibria (in symmetric profiles), and possibly more depending on the strength of preferences (the values of $\gamma_A$ and $\gamma_B$) and the structure of the network.

One goal of stochastic evolutionary game theory is to make predictions on which equilibrium is more likely to emerge as games are played repeatedly. This is the focus of the remainder of the paper, where the game I analyse is the Language Game. Given that strict equilibria are self enforcing, we will need an equilibrium selection device. The equilibrium selection criterion I use is stochastic stability (Foster and Young, 1990).

Let me first describe how population behaviour evolves. Time is discrete, and is indexed by $t = 0, 1, 2, \ldots$. Undiscounted utilities are received every period. Each period, one player is drawn randomly from the population and afforded the opportunity to update his/her current action. Thus the only feasible one-period transitions are to strategy profiles, that I now refer to as "states" in "state space" $\mathbf{S}$, that differ in a maximum of one component.

I suppose that the likelihood of playing a particular action is exponentially related to its expected payoff, where a player correctly forecasts the behaviour of others to remain unchanged. To be precise, if the current population profile is $\mathbf{s}$, and player $i$ is selected, then he/she chooses action $s_i \in S$ according to the probability distribution $p_i^\beta(s_i|\mathbf{s})$, where for any $\beta > 0$,

$$p_i^\beta(s_i|\mathbf{s}) \quad := \quad \frac{\exp\big(\beta\, U_i(s_i;\mathbf{s})\big)}{\exp\big(\beta\, U_i(a;\mathbf{s})\big) + \exp\big(\beta\, U_i(b;\mathbf{s})\big)} \tag{3.3}$$

As $\beta \to \infty$, player $i$'s response distribution as defined in (3.3) approaches that of a best response. This is true since actions that yield higher payoffs receive more weight as $\beta$ increases. So for any finite $\beta$, (3.3) is a perturbation of the best response. By combining the perturbed best response dynamics for each player, an irreducible Markov process with transition matrix $P^\beta$ is induced on the state space $\mathbf{S}$. The process is aperiodic and irreducible, and as such $P^\beta$ has a unique stationary distribution $\mu^\beta$.

The goal is to calculate the states to which the limiting distribution, as the perturbations of the process vanish, assign positive probability. To do this, we would like to know, $\mu^\star$, where

$$\mu^\star := \lim_{\beta\to\infty} \mu^\beta$$

We then have the following definition, originally due to Foster and Young (1990),

**Definition 20.** A state $\mathbf{s}$ is said to be *stochastically stable* if $\mu^\star(\mathbf{s}) > 0$, and *uniquely stochastically stable* if $\mu^\star(\mathbf{s}) = 1$.

Informally, whatever state(s) $\mu^\star$ assigns positive probability to, is (are) the predicted outcome(s) for the above system when it is constantly subjected to myopic best responses with noise.

For a given Language Game, $\mathcal{L}$, the set of stochastically stable states is denoted $\Xi(\mathcal{L})$. The approach to computing $\Xi(\mathcal{L})$ when errors are uniform, is to ignore computation of $\mu^\beta$, and instead to compute $\mu^\star$ directly using tools from the theory of dynamical systems (Freidlin and Wentzell, 1998) by analysing the recurrent classes of the unperturbed process $P := \lim_{\beta\to\infty} P^\beta$. When the errors are not uniform this is difficult, but fortunately, given the particular structure of the Language Game, there is a very workable shortcut such that these sophisticated techniques are not needed. The shortcut exploits the fact that the Language Game is a potential game (Shapley and Monderer, 1996). A game is said to be a potential game if the change in each player's utility from choosing a new action can be derived from a common function, referred to as the game's *potential function*. Formally,

**Definition 21.** An $N$-person game $G$ is a *potential game* if there exists a function $\rho : \mathbf{S} \to \mathbb{R}$ such that, for every player $i$, and every $\mathbf{s} \in \mathbf{S}$, and all pairs of actions $s_i, s_i' \in S_i$,

$$U_i(s_i; \mathbf{s}) - U_i(s_i'; \mathbf{s}) = \rho(s_i; \mathbf{s}) - \rho(s_i'; \mathbf{s}) \tag{3.4}$$

First I show that each local-interaction game of the Language Game, $\mathcal{L}$, is a potential game. This is easy. It can be checked that each of $G^{AA}, G^{AB}$, and $G^{BB}$ have potential given by[16] $\rho^{AA} : S \times S \to \mathbb{R}$, $\rho^{AB} : S \times S \to \mathbb{R}$, and $\rho^{BB} : S \times S \to \mathbb{R}$, where each is defined as follows (for $\rho^{AB}$, the first argument refers to the Group $A$ player's strategy, and the second to the Group $B$ player's strategy):

$$\begin{aligned}
\rho^{AA}(a,a) &= \gamma_A & \rho^{AA}(a,b) &= 0\\
\rho^{AA}(b,a) &= 0 & \rho^{AA}(b,b) &= 1-\gamma_A\\
\\
\rho^{AB}(a,a) &= \gamma_A + 1 - \gamma_B & \rho^{AB}(a,b) &= \gamma_A\\
\rho^{AB}(b,a) &= 1-\gamma_B & \rho^{AB}(b,b) &= 1 \qquad (3.5)\\
\\
\rho^{BB}(a,a) &= 1-\gamma_B & \rho^{BB}(a,b) &= 0\\
\rho^{BB}(b,a) &= 0 & \rho^{BB}(b,b) &= \gamma_B
\end{aligned}$$

Since each local-interaction of the Language Game is a potential game, by Theorem 2 of Neary (2010c), the Language Game is itself a potential game, with potential function, $\rho^\star$, where for any strategy profile $\mathbf{s}$,

$$\rho^\star(\mathbf{s}) := \sum_{\substack{(ij)\in E^{AA}\\ i<j}} \rho^{AA}(s_i, s_j) + \sum_{\substack{(ik)\in E^{AB}\\ i\in A,\, k\in B}} \rho^{AB}(s_i, s_k) + \sum_{\substack{(hk)\in E^{BB}\\ h<k}} \rho^{BB}(s_h, s_k) \tag{3.6}$$

[16] I should mention that the potential function for a game is never unique. For example, adding a constant will produce a new potential function.

The following is the key result. (It is a direct implication of Theorem 3 in Neary (2010c), which itself is an extension of a result from Blume (1993).)

**Theorem 10.** *For every $\beta > 0$, $P^\beta$ has the unique stationary distribution, $\mu^\beta$, given by*

$$\mu^\beta(\mathbf{s}) \quad = \quad \frac{exp\big(\gamma\rho^\star(\mathbf{s})\big)}{\sum_{\mathbf{s}'\in\mathbf{S}} exp\big(\gamma\rho^\star(\mathbf{s}')\big)} \tag{3.7}$$

*where $\rho^\star$ is the potential function of $\mathcal{L}$ as defined in equation (3.6).*
*Furthermore, $\Xi(\mathcal{L}) = \text{argmax}_{\mathbf{s}\in\mathbf{S}}\, \rho^\star(\mathbf{s})$.*

## 3.5 Selection

By Theorem 10 above, for a given Language Game, to compute the stochastically stable equilibria, what is required is to find those equilibrium profiles that maximize the potential function $\rho^\star$. This can be reduced to an integer programming problem as follows.

Without loss of generality, identify $\{a, b\}$ with $\{0, 1\}$. It is then the case that we can rewrite the potentials given in (3.5), where, for any $i, j \in A$ and $h, k \in B$,

$$\begin{aligned}
\rho^{AA}(s_i, s_j) &= (1 - s_i)(1 - s_j)\gamma_A + s_i s_j(1 - \gamma_A) \\
\rho^{AB}(s_j, s_k) &= (1 - s_j)(1 - s_k)(\gamma_A + 1 - \gamma_B) + (1 - s_j)s_k\gamma_A \\
&\quad + s_j(1 - s_k)(1 - \gamma_B) + s_j s_k 1 \\
\rho^{BB}(s_h, s_k) &= (1 - s_h)(1 - s_k)(1 - \gamma_B) + s_h s_k(\gamma_B)
\end{aligned}$$

Thus the problem of $\max_{\mathbf{s}\in\mathbf{S}} \rho^\star(\mathbf{s})$ reduces to

$$\max_{\{s_1,\ldots,s_N\}\in\{0,1\}^N} \left\{ \sum_{\substack{i,j\in A \\ i<j}} g_{ij}\, \rho^{AA}(s_i, s_j) + \sum_{i\in A, k\in B} g_{ik}\, \rho^{AB}(s_i, s_k) + \sum_{\substack{h,k\in B \\ h<k}} g_{hk}\, \rho^{BB}(s_h, s_k) \right\}$$

## 3.6 The Fully Connected Network

From Neary (2010a), when $\Gamma$ is fully connected there are only 3 candidate equilibrium profiles, $\{(\mathbf{a},\mathbf{a}),(\mathbf{a},\mathbf{b}),(\mathbf{b},\mathbf{b})\}$, where $(\mathbf{a},\mathbf{b})$ is the profile where all Group $A$ members adopt action $a$, while all those in Group $B$ choose action $b$, etc. Behaviour at these profiles is *group symmetric* in that all players in a given group take the the same action. Both $(\mathbf{a},\mathbf{a})$ and $(\mathbf{b},\mathbf{b})$ are always equilibria. Profile $(\mathbf{a},\mathbf{b})$ is an equilibrium provided the ratios of preferred payoff to less preferred payoff $\big(\gamma_A/(1-\gamma_A)$ for Group $A$ and $\gamma_B/(1-\gamma_B)$ for Group $B\big)$ are sufficiently high.[17]

Recall that for any two nonempty subsets of vertices, $X, Y \subseteq \mathcal{N}$, $E^{XY}$ is the set of all undirected edges $(i,j)$ with $i \in X$ and $j \in Y$. Let $e(X,Y) = |E^{XY}|$ be the number of such edges.

I now compute the potential for each candidate group symmetric profile. To do so, the following facts will be useful. The number of edges on any fully connected undirected graph with $N$ vertices is $\binom{N}{2}$. So, $e(A,A) = \binom{N^A}{2}$, $e(B,B) = \binom{N^B}{2}$. Furthermore, $\Gamma[E^{AB}] = (A, B, E^{AB})$ is a fully connected bipartite graph, so that $e(A,B) = N^A \times N^B$.

We have that,

$$
\begin{array}{lccccc}
 & E^{AA} & & E^{AB} & & E^{BB} \\[2ex]
\rho^\star\big((\mathbf{a},\mathbf{a})\big) = & \binom{N^A}{2}\cdot\gamma_A & + & (N^A N^B)\cdot(\gamma_A + 1 - \gamma_B) & + & \binom{N^B}{2}\cdot(1-\gamma_B) \\[2ex]
\rho^\star\big((\mathbf{a},\mathbf{b})\big) = & \binom{N^A}{2}\cdot\gamma_A & + & (N^A N^B)\cdot\gamma_A & + & \binom{N^B}{2}\cdot\gamma_B \\[2ex]
\rho^\star\big((\mathbf{b},\mathbf{b})\big) = & \binom{N^A}{2}\cdot(1-\gamma_A) & + & (N^A N^B)\cdot 1 & + & \binom{N^B}{2}\cdot\gamma_B
\end{array} \tag{3.8}
$$

where the first term on the right hand side is the potential of edges $E^{AA}$, etc.

The stochastically stable equilibria are those for which $\rho^\star$ takes the largest

[17]Precisely, it requires: $(N^A - 1)\gamma_A \geq N^B(1-\gamma_A)$ and $(N^B - 1)\gamma_B \geq N^A(1-\gamma_B)$. If one group is strictly larger, then the relevant inequality always holds, and so one must only check that of the smaller group.

value out of the three above. Simple algebra shows that for $(\mathbf{a},\mathbf{b})$ to be stochastically stable, it must be that

$$\gamma_A \geq \gamma_A^\star = \frac{1}{2} + \frac{1}{2}\frac{N^B}{N-1} \text{ and } \gamma_B \geq \gamma_B^\star = \frac{1}{2} + \frac{1}{2}\frac{N^A}{N-1} \tag{3.9}$$

From the omitted algebra, it is the condition $\rho^\star\big((\mathbf{a},\mathbf{b})\big) \geq \rho^\star\big((\mathbf{a},\mathbf{a})\big)$ that gives the lower bound of $\gamma_B^\star$. It is intuitive that comparison of these two states yield a requirement for a Group $B$ player's payoff, since obviously all Group $A$ players prefer $(\mathbf{a},\mathbf{a})$.

Footnote 17 states that for $(\mathbf{a},\mathbf{b})$ to be an equilibrium, it must be that $(N^B-1)\gamma_B \geq N^A(1-\gamma_B)$, which happens if and only if $\gamma_B \geq \bar{\gamma}_B = N^A/(N-1)$. It is clear that both $\gamma_A^\star \geq \bar{\gamma}_A$ and $\gamma_B^\star \geq \bar{\gamma}_B$, so that a necessary condition for profile $(\mathbf{a},\mathbf{b})$ to be stochastically stable is that it be an equilibrium.

Here is the main result concerning equilibrium selection for the fully connected language game.

**Theorem 11.** *For the Language Game, $\mathcal{L}$, with $\Gamma$ fully connected,*

- $(\mathbf{a},\mathbf{a}) \in \Xi(\mathcal{L})$, *if and only if*

$$\binom{N^A}{2}(2\gamma_A - 1) - \binom{N^B}{2}(2\gamma_B - 1) + N^A N^B(\gamma_A - \gamma_B) \geq 0 \quad \textit{and} \quad \gamma_B \leq \gamma_B^\star \tag{3.10}$$

- $(\mathbf{a},\mathbf{b}) \in \Xi(\mathcal{L})$, *if and only if*

$$\gamma_A \geq \gamma_A^\star \quad \textit{and} \quad \gamma_B \geq \gamma_B^\star \tag{3.11}$$

- $(\mathbf{b},\mathbf{b}) \in \Xi(\mathcal{L})$, *if and only if*

$$\binom{N^A}{2}(2\gamma_A - 1) - \binom{N^B}{2}(2\gamma_B - 1) + N^A N^B(\gamma_A - \gamma_B) \leq 0 \quad \textit{and} \quad \gamma_A \leq \gamma_A^\star \tag{3.12}$$

*Proof.* The proof follows by combining the first and third equations of (3.8) with the inequalities in (3.9). □

Theorem 11 demonstrate thes simplicity of the potential approach. Basically, one just computes the number of edges in the fully connected network and weights them according to the local-interaction potentials.

Consider conditions (3.10) and (3.12) of Theorem 11. In (3.10) for example, the first inequality compares the potential at $(\mathbf{a},\mathbf{a})$ to that at $(\mathbf{b},\mathbf{b})$, while the second compares that at $(\mathbf{a},\mathbf{a})$ to that at $(\mathbf{a},\mathbf{b})$. The second inequality is just the reverse of that in (3.9), which now requires that Group $B$ players do not have overly strong preferences. The first inequality computes the number of edges in the fully connected network and weights them according to the local-interaction potentials at each of the symmetric population profiles.

For the sake of intuition, the following two special cases of Theorem 11 are useful. First imagine that both within-group local-interaction are payoff mirror image. That is, $\gamma_A = \gamma_B$. In this situation it can be shown that,

$$\rho^\star\big((\mathbf{a},\mathbf{a})\big) \geq \rho^\star\big((\mathbf{b},\mathbf{b})\big) \iff (2\gamma_A - 1)\Big[\binom{N^A}{2} - \binom{N^B}{2}\Big] \geq 0$$

which, given that $2\gamma_A > 1$, happens only if $\binom{N^A}{2} \geq \binom{N^B}{2}$, which happens only if $N^A > N^B$ - i.e. Group $A$ is the bigger group.

Now, another useful exercise is to examine what happens when within-group local-interactions are not payoff mirror image, but each group has the same number of members. That is, when $N^A = N^B$. In this case, it can easily be shown that

$$\text{sign}\Big[\rho^\star\big((\mathbf{a},\mathbf{a})\big) - \rho^\star\big((\mathbf{b},\mathbf{b})\big)\Big] = \text{sign}\big[\gamma_A - \gamma_B\big] \tag{3.13}$$

Equation 3.13 above can be used to indicate the fragility of the stochastic stability prediction in the payoff dependent case. Recall back to the example in Section 4.2, when the population was fully connected (Figures 3.1 and 3.2). I stated that the uniform error dynamics of KMR and the logit error dynamics select both symmetric equilibria each with probability 0.5. However, if we perturb the payoff of the most desired outcome for Group $A$, from 3 to $3+\varepsilon$, for some $\varepsilon > 0$, then under the uniform errors of KMR, the prediction does not change. But, under the logit errors which measures the potential at different profiles, it is clear from (3.5) that everybody adopting action $a$ is uniquely stochastically stable. This selection result holds true even for the subgraph $\hat{\Gamma}$ in Figure 3.2.

I have just proved the following two Theorems, which are really just special cases of Theorem 11

**Theorem 12.** *Suppose within-group local-interactions of the fully connected Language Game are payoff mirror image. Then,*

- *If $N^A < N^B$, then*

$$\Xi(\mathcal{L}) \ni \begin{cases} (\mathbf{a}, \mathbf{b}), & \text{if } \gamma_A \geq \gamma_A^\star, \\ (\mathbf{b}, \mathbf{b}), & \text{if } \gamma_A \leq \gamma_A^\star, \end{cases}$$

- *If $N^A > N^B$, then*

$$\Xi(\mathcal{L}) \ni \begin{cases} (\mathbf{a}, \mathbf{a}), & \text{if } \gamma_B \leq \gamma_B^\star, \\ (\mathbf{a}, \mathbf{b}), & \text{if } \gamma_B \geq \gamma_B^\star, \end{cases}$$

**Theorem 13.** *Consider the fully connected Language Game with $N^A = N^B$. Then*

$$\Xi(\mathcal{L}) \ni \begin{cases} (\mathbf{a}, \mathbf{a}), & \text{if } \gamma_A > \gamma_B \geq \gamma_B^\star, \\ (\mathbf{a}, \mathbf{b}), & \text{if } \gamma_A \geq \gamma_A^\star \text{ and } \gamma_B \geq \gamma_B^\star, \\ (\mathbf{b}, \mathbf{b}), & \text{if } \gamma_B > \gamma_A \geq \gamma_A^\star, \end{cases}$$

Theorem 12 says that when payoffs are mirrored, if one group has strictly more members than the other, the stochastically stable equilibrium cannot involve all players uniformly adopting the preferred outcome of the smaller group.

Theorem 13 says that is both groups are of equal size the long run outcome will involve all players uniformly adopting the preferred outcome of the group with stronger preferences, unless both $\gamma_A \geq \gamma_A^\star$ and $\gamma_B \geq \gamma_B^\star$.

Theorems 11 - 13 provide somewhat different selection results to those of Neary (2010c) who uses the uniform errors framework of KMR. It is of some interest to compare the two.

## 3.7 Some Commonly Studied Networks

Let us quickly examine which states maximize the potential function for some networks that have received a lot of attention in the literature.

**Graph 1** (2 Star Networks with centers connected)**.** Each group is arranged in a star network. The central player for Group $A$ ($B$) is labelled $A_1$ ($B_1$). Any player who is not central is referred to as *peripheral*. Thus, for $\Gamma[A]$, $g_{A_1 A_j} = 1$ if and only if $j \in \{2, \ldots, N^A\}$, while for $\Gamma[B]$, $g_{B_1 B_k} = 1$ if and only if $k \in \{2, \ldots, N^B\}$. I also assume that the central players of each star are adjacent ($g_{A_1 B_1} = 1$), and that there are no other links. Clearly, in any pure strategy equilibrium, the peripheral players must take the same action as the central player of their Group.[18] It can quite easily be seen that the only candidate equilibria to this game are $\{(\mathbf{a}, \mathbf{a}), (\mathbf{a}, \mathbf{b}), (\mathbf{b}, \mathbf{a}), (\mathbf{b}, \mathbf{b})\}$, where $(\mathbf{b}, \mathbf{a})$ is an equilibrium only if $(N^A - 1)(1 - \gamma_A) \geq \gamma_A$ and $(N^B - 1)(1 - \gamma_B) \geq \gamma_B$.

Figure 1 below illustrates this network. The left star represents $\Gamma[A]$, and the right star $\Gamma[B]$. Group $A$ members are red vertices, while Group $B$ members are blue vertices. The central player of each subnetwork is depicted larger than the others. I assume that $N^A > N^B$.

Let us compute the potential for each of these equilibrium profiles. Note that $e(A, A) = N^A - 1$, $e(A, B) = 1$, and $e(B, B) = N^B - 1$.

$$\begin{array}{rcccc}
\rho^\star\big((\mathbf{a}, \mathbf{a})\big) = & (N^A - 1) \cdot \gamma_A & + & 1 \cdot (\gamma_A + 1 - \gamma_B) & +(N^B - 1) \cdot (1 - \gamma_B) \\
\rho^\star\big((\mathbf{a}, \mathbf{b})\big) = & (N^A - 1) \cdot \gamma_A & + & 1 \cdot \gamma_A & +(N^B - 1) \cdot \gamma_B \\
\rho^\star\big((\mathbf{b}, \mathbf{a})\big) = & (N^A - 1) \cdot (1 - \gamma_A) & + & 1 \cdot (1 - \gamma_B) & +(N^B - 1) \cdot (1 - \gamma_B) \\
\rho^\star\big((\mathbf{b}, \mathbf{b})\big) = & (N^A - 1) \cdot (1 - \gamma_A) & + & 1 \cdot 1 & +(N^B - 1) \cdot \gamma_B
\end{array}$$

It is clear by inspection that $\rho^\star\big((\mathbf{b}, \mathbf{a})\big)$ is the smallest value, and thus $\big((\mathbf{b}, \mathbf{a})\big)$ can never be stochastically stable. This is interesting as this profile has the lowest welfare properties of all equilibria, as it has each group coordinating on the other group's preferred action.

I now compare the potential at each of the remaining equilibria. The fol-

[18] In fact this game has no equilibria in totally mixed strategies.

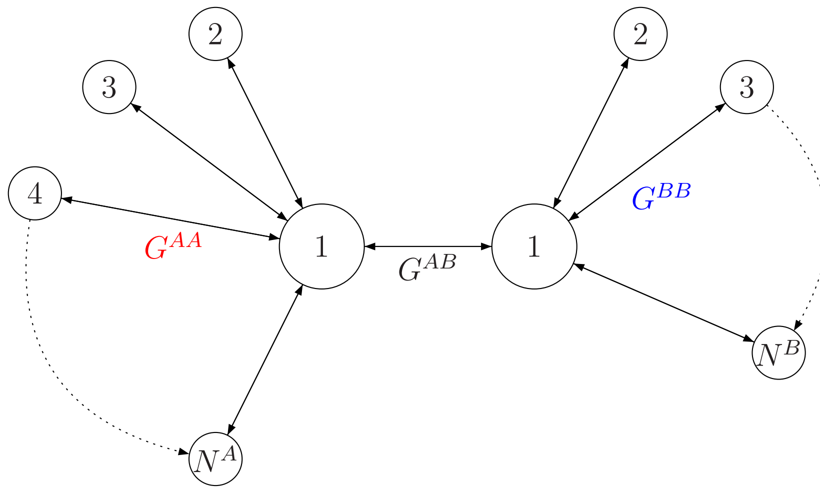


**Figure 3.4**: Two stars connected center to center

lowing are the relevant inequalities,

$$\rho^\star\big((\mathbf{a},\mathbf{b})\big) \geq \rho^\star\big((\mathbf{a},\mathbf{a})\big) \iff \gamma_B \geq \frac{N^B}{2N^B-1} \tag{3.14}$$

$$\rho^\star\big((\mathbf{a},\mathbf{b})\big) \geq \rho^\star\big((\mathbf{b},\mathbf{b})\big) \iff \gamma_A \geq \frac{N^A}{2N^A-1} \tag{3.15}$$

Note that the utility to the central Group $B$ player, $B_1$, of choosing action $b$ at state $(\mathbf{a},\mathbf{b})$ is $(N^B-1)\gamma_B$, while that of choosing action $a$ is $N^B(1-\gamma_B)$. Note that by comparing these two, the second inequality in (3.14) falls out, and thus it really captures the decision of $B_1$.

Given that both $\gamma_A, \gamma_B > 1/2$, the right hand side inequalities of (3.14) and (3.15) can always hold for sufficiently large values of $N^A$ and $N^B$. Thus, perhaps unsurprisingly, the most likely outcome for two star networks connected center to center is for each star to coordinate on its most preferred action. Note that this will be true once each group reaches a given absolute size, and is independent of the relative size of the groups.

This example and prediction can be contrasted with the stochastic stability results for a homogeneous group arranged in a star network. Jackson and Watts (2002, 2008) show that under the uniform error dynamics, that either symmetric profile is stochastically stable (the intuition being that all it takes to move between equilibria is for the central player to mutate, which he/she will do with equal likelihood). With logit error dynamics, risk dominance will be selected.

The analysis of uniform error dynamics is similarly easy for this situation. In fact, any group symmetric profile that is an equilibrium is stochastically stable according to such dynamics, since all it takes to transition between them is for a central player to mutate (which, again, he/she will do with equal likelihood). However, we just saw that the profile $(\mathbf{a}, \mathbf{b})$ will be uniquely stochastically stable for large enough group sizes under the payoff dependent dynamics.

**Graph 2** (2 Star Networks connected spoke to spoke)**.** Again, suppose that each group is arranged in a star network. Again, $A_1$ and $B_1$ are the respective central players, and $N^A > N^B$. This time however, assume that each peripheral agent from Group $B$ is connected to the central Group $B$ player and also to a unique peripheral player from Group $A$. The set of links, $E$, can thus be defined by: $g_{A_i A_j} = 1$ if and only if $i = 1$ and $j \in \{2, \ldots, N^A\}$; $g_{B_k B_l} = 1$ if and only if $k = 1$ and $l \in \{2, \ldots N^B\}$. And finally $g_{A_m B_n} = 1$ if and only if $m = n \in \{2, \ldots, N^B\}$. Figure 2 illustrates this, with group colours as before. Note that since $N^A > N^B$, there are $(N^A - 1) - (N^B - 1) = N^A - N^B$ "left over" peripheral Group $A$ players who are only connected to the central Group $A$ player, $A_1$.

To help compute the pure strategy equilibria, note that in any equilibrium, the "left over" peripheral players in Group $A$, $\{A_{N^B+1}, \ldots, A_{N^A}\}$, must coordinate with $A_1$. All other peripheral players will only choose their least preferred equilibrium action only if both their neighbours are using it.

Clearly $(\mathbf{a}, \mathbf{a})$ and $(\mathbf{b}, \mathbf{b})$ are equilibria. It is easily checked that $(\mathbf{a}, \mathbf{b})$ is an equilibrium, $(\mathbf{b}, \mathbf{a})$ is not, and it that, $\mathbf{s} = \big((s_{A_1}, s_{A_2}, \ldots, s_{A_{N^B}}, s_{A_{N^B+1}}, \ldots, s_{A_{N^A}}), \mathbf{s}_B\big) = \big((b, a, \ldots, a, b, \ldots, b), \mathbf{a}\big)$ is an equilibrium if and only if the following (which is a requirement on $A_1$) holds: $(N^A - 1 - N^B)(1 - \gamma_A) > (N^B - 1)\gamma_A$. However, it is always the case that $\rho^\star\big((\mathbf{b}, \mathbf{b})\big) > \rho^\star\big((b, a, \ldots, a, b, \ldots, b), \mathbf{a}\big)$.

Given this, $(\mathbf{a}, \mathbf{a}), (\mathbf{a}, \mathbf{b})$, and $(\mathbf{b}, \mathbf{b})$ are the only candidates for stochastic stability. The potential at each of these equilibrium profiles is,

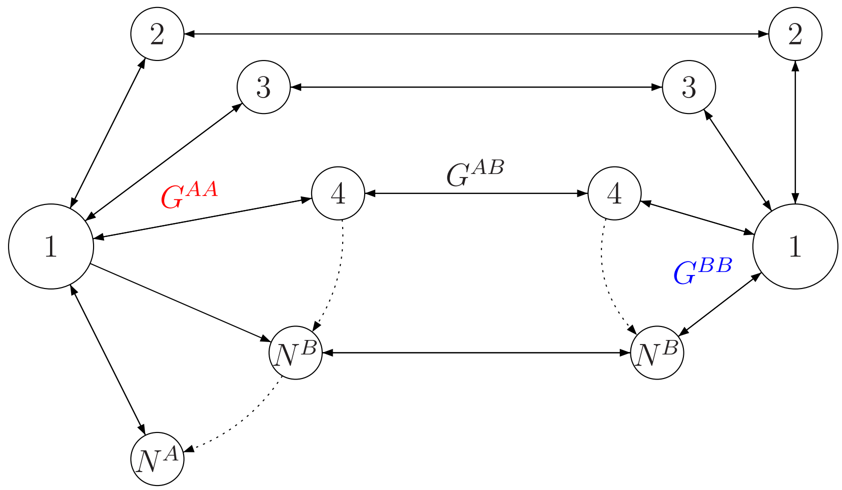


**Figure 3.5**: Two stars connected spokes to spokes.

$$\begin{aligned}
\rho^\star\big((\mathbf{a},\mathbf{a})\big) &= (N^A-1)\gamma_A + (N^B-1)(\gamma_A+1-\gamma_B) &+& \quad (N^B-1)(1-\gamma_B)\\
\rho^\star\big((\mathbf{a},\mathbf{b})\big) &= (N^A-1)\gamma_A + (N^B-1)\gamma_A &+& \quad (N^B-1)\gamma_B\\
\rho^\star\big((\mathbf{b},\mathbf{b})\big) &= (N^A-1)(1-\gamma_A) + (N^B-1)1 &+& \quad (N^B-1)\gamma_B
\end{aligned}$$

The following are then easily shown:

$$\rho^\star\big((\mathbf{a},\mathbf{a})\big) \geq \rho^\star\big((\mathbf{a},\mathbf{b})\big) \iff 2/3 \geq \gamma_B$$

$$\rho^\star\big((\mathbf{a},\mathbf{a})\big) \geq \rho^\star\big((\mathbf{b},\mathbf{b})\big) \iff$$

$$(N^A-1)(2\gamma_A-1) + (N^B-1)\gamma_A \geq (N^B-1)(3\gamma_B-1)$$

$$\rho^\star\big((\mathbf{b},\mathbf{b})\big) \geq \rho^\star\big((\mathbf{a},\mathbf{b})\big) \iff (N^A+N^B-2) \geq \gamma_A(2N^A+N^B-3)$$

An interesting pattern which has emerged in all of the examples is that the number of links in each group's subnetwork is important as regards selection. When the within-group games are payoff mirror image, this becomes hugely important. This is summarized in the following result, again stated without proof.

**Theorem 14.** *Consider the Language Game, $\mathcal{L}$, where local-interactions are payoff mirror image. The following always holds,*

- *If $e(A,A)<e(B,B)$, then $(\mathbf{a},\mathbf{a})$ is not stochastically stable.*
- *If $e(A,A)>e(B,B)$, then $(\mathbf{b},\mathbf{b})$ is not stochastically stable.*

Theorem 14 above is nothing more than statement of the following identity: $\text{sign}\Big(e(A,A) - e(B,B)\Big) = \text{sign}\Big(\rho^\star\big((\mathbf{a},\mathbf{a})\big) - \rho^\star\big((\mathbf{b},\mathbf{b})\big)\Big)$. I should emphasize that this corollary applies to the number of links within-group, and not to the number of members in each group. While the number of agents within a given group does bound from above the number of possible links within-group, it is the number of links themselves that is the important feature.

It is important to note however, that the above does not say that if $e(A,A) < e(B,B)$, then $(\mathbf{b},\mathbf{b})$ will be stochastically stable. Nor does it say that all Group $A$ members taking action $a$ will be not be part of the stochastically stable outcome. Rather it just says that in the long run, whichever group has more within-group connections must contain some subset who take their preferred action.

A similar result can be shown in the situation where the within-group local-interaction are not payoff mirror image.

**Theorem 15.** *Consider the Language Game, $\mathcal{L}$.*

- *If $\gamma_A > \gamma_B$, then*

$$e(A,A) \geq e(B,B) \Rightarrow \rho^\star\big((\mathbf{a},\mathbf{a})\big) > \rho^\star\big((\mathbf{b},\mathbf{b})\big)$$

- *If $\gamma_A < \gamma_B$, then*

$$e(A,A) \leq e(B,B) \Rightarrow \rho^\star\big((\mathbf{a},\mathbf{a})\big) < \rho^\star\big((\mathbf{b},\mathbf{b})\big)$$

The following example helps illustrate the findings.

**Graph 3.** There are twelve players, Group $A = \{A_1, \ldots, A_4\}$ and Group $B = \{B_5, \ldots, B_{12}\}$. Suppose Group $A$ are tightly knit, such that $\Gamma[A]$ is fully connected. Group $B$ is arranged in a line, $E^{BB} = \{(i, i+1)\}_{i=4,\ldots,11}$. All Group $A$ players are connected to players 5 and 6, and that there are no other edges in $\Gamma[E^{AB}]$.

The network $\Gamma$ is displayed in Figure 3.6 below. Group $A$ members are contained in the red box that is shaded.[19] The dashed lines connecting Group $B$ players, $B_5$ and $B_6$, to Group $A$ indicates that $g_{A_iB_5} = g_{A_iB_6} = 1$ for all $i = 1, \ldots, 4$.

[19]This is to convey that all pairs of them interact without having to draw the many connections.

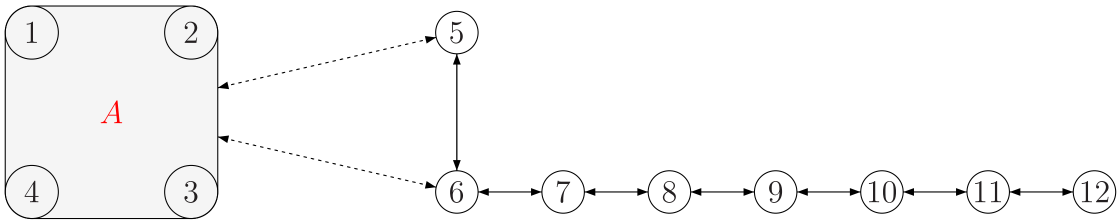


**Figure 3.6**: $\Gamma$

Simple arithmetic yields that $e(B,B) = 6$, $e(A,B) = 8$, and $e(B,B) = 7$. Since $e(B,B) > e(B,B)$, the corollary to Theorem 14 states that if $\gamma_A = \gamma_B$, then in the long run there is some nonempty subset of Group $B$ that take action $b$.

Let us now analyse this game when $\gamma_A = \gamma_B$ for a variety of different values. In each case, I will first write down the set of pure strategy equilibria and then compute the stochastically stable outcome by examining the potential of each. The set of candidate equilibria is always the same: $\Big\{(\mathbf{a},\mathbf{a}), \big(\mathbf{a},(a,a,b,\ldots,b)\big), (\mathbf{a},\mathbf{b}), (\mathbf{b},\mathbf{b})\Big\}$

The potential for each of these profiles is given below,

$$\begin{array}{lclclclcl}
\rho^\star\big(E_1\big) & = & 6\cdot q & + & 8\cdot(p+q) & + & 7\cdot p \\
\rho^\star\big(E_2\big) & = & 6\cdot p & + & 8\cdot p & + & 7\cdot p \\
\rho^\star\big(E_3\big) & = & 6\cdot p & + & 8\cdot(p+q) & + & 5\cdot p \\
\rho^\star\big(E_4\big) & = & 6\cdot p & + & 8\cdot(p+q) & + & 7\cdot q
\end{array}$$

**Example 7** ($\alpha = 4/7$)**.** In this case the equilibrium set is $\Big\{E_1, E_3, E_4\Big\}$. We have that

$$\Big\{\rho^\star\big(E_1\big), \rho^\star\big(E_2\big), \rho^\star\big(E_3\big), \rho^\star\big(E_4\big)\Big\} = \Big\{102, 84, 100, 101\Big\}$$

Clearly then the stochastically stable outcome is $E_1 = (\mathbf{b},\mathbf{b})$.

**Example 8** (($\alpha = 5/8$)**.** In this case the equilibrium set is $\Big\{E_1, E_3, E_4\Big\}$. We have that

$$\Big\{\rho^\star\big(E_1\big), \rho^\star\big(E_2\big), \rho^\star\big(E_3\big), \rho^\star\big(E_4\big)\Big\} = \Big\{117, 105, 119, 115\Big\}$$

Clearly then the stochastically stable outcome is $E_3 = \big(\mathbf{a},(a,a,b,\ldots,b)\big)$.

**Example 9** ($\alpha = 5/7$)**.** In this case the equilibrium set is $\Big\{E_1, E_3, E_4\Big\}$. We have that

$$\Big\{\rho^\star\big(E_1\big), \rho^\star\big(E_2\big), \rho^\star\big(E_3\big), \rho^\star\big(E_4\big)\Big\} = \Big\{103, 105, 111, 100\Big\}$$

Clearly then the stochastically stable outcome is $E_3 = \big(\mathbf{a}, (a, a, b, \ldots, b)\big)$.

**Example 10** ($\alpha = 7/8$)**.** In this case the equilibrium set is $\Big\{E_1, E_2, E_4\Big\}$. We have that

$$\Big\{\rho^\star\big(E_1\big), \rho^\star\big(E_2\big), \rho^\star\big(E_3\big), \rho^\star\big(E_4\big)\Big\} = \Big\{119, 147, 141, 113\Big\}$$

Clearly then the stochastically stable outcome is $E_2 = (\mathbf{a}, \mathbf{b})$.

The above examples correspond to the following ranges on $\alpha$. If $\alpha \in (1/2, 3/5)$, then $(\mathbf{b}, \mathbf{b})$ is uniquely stochastically stable. If $\alpha \in (3/5, 4/5)$ then $\big(\mathbf{a}, (a, a, b, \ldots, b)\big)$ is uniquely stochastically stable. While if $\alpha \in (4/5, 1)$ then $(\mathbf{a}, \mathbf{b})$ is uniquely stochastically stable.

The intuition is quite straightforward. When $\alpha$ is sufficiently small, then no player minds too much coordinating on their less preferred outcome. Hence the greater number of connections in Group $B$ dominate. When $\alpha$ is not to small nor too large, then players strongly prefer to coordinate with their neighbours on their most preferred action but can be persuaded to choose their least preferred action provided enough of their neighbours are choosing it. This is relevant for players $B_5$ and $B_6$. When $\alpha$ is sufficiently large then provided at least one neighbour is taking your preferred action then it is worth sacrificing all other interactions for this one successful coordination.

## 3.8 Failure to Coordinate Across Group

The linear program in Section 3.5 computes the stochastically stable equilibria for the Language Game on any arbitrary network. The purpose of this Section is to determine what conditions must be placed on the subnetworks, $\Gamma[A], \Gamma[B]$, and $\Gamma[E^{AB}]$, to determine when state $(\mathbf{a}, \mathbf{b})$ is stochastically stable.

It turns out that such a characterization is very straightforward. Theorem

16 below provides straightforward sufficient conditions.[20]

**Theorem 16.** *The bipartite graph* $\Gamma[E^{AB}]$ *is the triple* $(V^A, V^B, E^{AB})$. *Let* $Q$ *denote the set of players in Group* $A$ *who use action* $b$*, and* $R$ *denote the set of players in Group* $B$ *who use action* $a$*. Sufficient conditions for* $(\mathbf{a}, \mathbf{b}) \in \Xi(\mathcal{L})$*, are that the following both hold*

$$\max_{\{Q:\, \emptyset \subset Q \subseteq A\}} \frac{e(Q,Q) + e(Q,V^B)}{e(Q,Q) + e(A \backslash Q, Q)} \quad \leq \quad \frac{\gamma_A}{1-\gamma_A} \tag{3.16}$$

$$\max_{\{R:\, \emptyset \subset R \subseteq B\}} \frac{e(R,R) + e(R,V^A)}{e(R,R) + e(B \backslash R, R)} \quad \leq \quad \frac{\gamma_B}{1-\gamma_B} \tag{3.17}$$

.

*Proof.* I show only the first inequality. The second is shown analogously.

Let $\mathbf{\Lambda}_A \subset \mathbf{S}$ denote the set of states where all Group $B$ players are constrained to take action $b$. Consider the discrete time process where players in $A$ update their action according to (3.3). Let $P^{\beta A}$ denote this irreducible process on state space $\mathbf{\Lambda}_A$ with states $\lambda$. As in Theorem 10, it can be shown that the restricted process $P^{\beta A}$ has unique stationary distribution, $\mu^{\beta A}$, where

$$\mu^{\beta A}(\boldsymbol{\lambda}) = \frac{\exp\big(\beta \rho^\star(\boldsymbol{\lambda})\big)}{\sum_{\boldsymbol{\lambda}' \in \mathbf{\Lambda}_A} \exp\big(\beta \rho^\star(\boldsymbol{\lambda}')\big)}$$

I will show that $\rho^\star$ is maximized at $(\mathbf{a}, \mathbf{b})$, provided the condition in the statement of the theorem holds. Fix a state $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_N) \in \mathbf{\Lambda}_A$, and note that

$$\rho^\star(\boldsymbol{\lambda}) = \sum_{(i,j) \in E^{AA}} \rho^{AA}(\lambda_i, \lambda_j) \;+\; \sum_{(h,k) \in E^{AB}} \rho^{AB}(\lambda_h, \lambda_k) \;+\; e(B,B)\rho^{BB}(b,b)$$

For any two sets $X, Y$, define $X \backslash Y := X \cap Y^c$, where $Y^c$ denotes the

[20]This result does not depend on the path-connectedness of $\Gamma[A]$ and $\Gamma[B]$.

complement of $Y$. Now letting $Q' = \{i \in A : \lambda_i = b\}$, it is easy to see that,

$$\begin{aligned}\rho^\star(\boldsymbol{\lambda}) &= \rho^{AA}(a,a)\cdot e(A\backslash Q', A\backslash Q') + \rho^{AA}(b,b)\cdot e(Q',Q') + \rho^{AB}(b,b)\cdot e(Q',V^B)\\ &+ \rho^{AB}(a,b)\cdot e(A\backslash Q', V^B) + \rho^{BB}(b,b)\cdot e(B,B)\\ &= \gamma_A\cdot e(A\backslash Q', A\backslash Q') + (1-\gamma_A)\cdot e(Q',Q') + 1\cdot e(Q',V^B)\\ &+ \gamma_A\cdot e(A\backslash Q', V^B) + \gamma_B\cdot e(B,B)\end{aligned}$$

and that

$$\begin{aligned}\rho^\star\big((\mathbf{a},\mathbf{b})\big) &= \rho^{AA}(a,a)\cdot e(A,A) + \rho^{AB}(a,b)\cdot e(A,V^B) + \rho^{BB}(b,b)\cdot e(B,B)\\ &= \gamma_A\cdot e(A,A) + \gamma_A\cdot e(A,V^B) + \gamma_B\cdot e(B,B)\end{aligned}$$

Thus

$$\begin{aligned}\rho^\star\big((\mathbf{a},\mathbf{b})\big) - \rho^\star(\boldsymbol{\lambda}) &= \gamma_A e(A,A) + \gamma_A e(A,V^B) - \gamma_A e(A\backslash Q', A\backslash Q')\\ &- (1-\alpha)e(Q',Q') - 1e(Q',V^B) - \alpha e(A\backslash Q', V^B)\\ &= \gamma_A\big[e(A,A) - e(A\backslash Q', A\backslash Q') + e(A,V^B) - e(A\backslash Q', V^B)\big]\\ &- \big[(1-\gamma_A)\cdot e(Q',Q') + e(Q',V^B)\big]\\ &= \gamma_A\big[e(A,Q') + e(A,V^B) - e(A\backslash Q', V^B)\big]\\ &- \big[(1-\gamma_A)\cdot e(Q',Q') + e(Q',V^B)\big]\\ &= \gamma_A\big[e(Q',A) + e(Q',V^B)\big]\\ &- \big[(1-\gamma_A)\cdot e(Q',Q') + e(Q',V^B)\big]\\ &= \gamma_A\big[e(Q',Q') + e(Q',A\backslash Q') + e(Q',V^B)\big]\\ &- \big[(1-\gamma_A)\cdot e(Q',Q') + e(Q',V^B)\big]\end{aligned}$$

We wish to find conditions on $Q'$ such that the above is positive. If $Q' = \emptyset$, then $\boldsymbol{\lambda} = (\mathbf{a},\mathbf{b})$ and the above expression is zero. So we restrict attention to nonempty $Q'$. Attention can also be restricted to sets $Q'$ such that $Q' \cap V^A \neq \emptyset$, since otherwise $e(V^B, Q') = 0$ and all the terms in the last line above are positive meaning the desired inequality follows trivially. Thus it is sufficient that

$$\frac{\gamma_A}{1-\gamma_A} \geq \frac{e(Q',Q') + e(Q',V^B)}{e(Q',Q') + e(A\backslash Q',Q')}$$

holds for all nonempty $Q' \subseteq A$ with $Q' \cap V^A \neq \emptyset$.

That we can make the stronger restriction that $Q' \subseteq V^A$, rather than $Q' \cap V^A \neq \emptyset$, follows by noting that

- If the inequality in (3.16) holds for $Q \subseteq V^A$, then it holds for any $\hat{Q} \subseteq A$ where $\hat{Q} \cap V^A = Q$.

- If the inequality in (3.17) holds for $R \subseteq V^B$, then it holds for any $\hat{R} \subseteq A$ where $\hat{R} \cap V^A = R$.

  To prove this, let $\boldsymbol{\lambda}$ denote the state in which agents $Q \subseteq V^A$ take action $b$. Specifically $Q = \left\{i \in V^A | \lambda_i = b\right\}$. Now consider a state $\hat{\boldsymbol{\lambda}}$ where $\hat{\lambda}_i = b \Rightarrow i \in \hat{Q}$, with $Q \subseteq \hat{Q} \subset A$ and $\hat{Q} \cap V^A = Q$.

  It is sufficient to show that

$$\rho^\star(\boldsymbol{\lambda}) - \rho^\star(\hat{\boldsymbol{\lambda}}) \geq 0$$

  Using Theorem 16 above we have that

$$\begin{aligned}
\rho^\star\big((\mathbf{a},\mathbf{b})\big) - \rho^\star(\hat{\boldsymbol{\lambda}}) \;\; = \;\; & \alpha\big[e(\hat{Q},A) + e(\hat{Q},V^B)\big] \\
& - \big[(1-\alpha)\cdot e(\hat{Q},\hat{Q}) + e(\hat{Q},V^B)\big] \\
& = \alpha \cdot e(\hat{Q},A) - (1-\alpha)\cdot e(\hat{Q},\hat{Q}) \\
& + \alpha \cdot e(Q,V^B) - e(Q,V^B)
\end{aligned}$$

So we have that

$$\begin{aligned}
\rho^\star(\boldsymbol{\lambda}) - \rho^\star(\hat{\boldsymbol{\lambda}}) &= \big[\rho^\star\big((\mathbf{a},\mathbf{b})\big) - \rho^\star(\hat{\boldsymbol{\lambda}})\big] - \big[\rho^\star\big((\mathbf{a},\mathbf{b})\big) - \rho^\star(\boldsymbol{\lambda})\big] \\
&= \big[\alpha \cdot e(\hat{Q}, A) - (1-\alpha)\cdot e(\hat{Q},\hat{Q})\big] \\
&- \big[\alpha \cdot e(Q, A) - (1-\alpha)\cdot e(Q,Q)\big] \\
&= \alpha \cdot e(\hat{Q}\backslash Q, A) - (1-\alpha)\cdot e(\hat{Q},\hat{Q}) + q \cdot e(Q,Q) \\
&= \alpha \cdot e(\hat{Q}\backslash Q, A) \\
&- (1-\alpha)\cdot \big[e(Q,Q) + e(\hat{Q}\backslash Q, \hat{Q}\backslash Q) + \cdot e(\hat{Q}\backslash Q, Q)\big] \\
&+ (1-\alpha)\cdot e(Q,Q) \\
&= \alpha \cdot e(\hat{Q}\backslash Q, A) - (1-\alpha)\cdot e(\hat{Q}\backslash Q, \hat{Q}) \\
&> 0
\end{aligned}$$

where the last inequality follows since $\alpha > 1/2$, and $e(\hat{Q}\backslash Q, A) \geq \cdot e(\hat{Q}\backslash Q, \hat{Q})$ due to the fact that $\hat{Q} \subseteq A$.

□

Consider the weak inequalities (3.16) and (3.17) in Theorem 16 above. Generically $\gamma_A$ and $\gamma_B$ are irrational, and hence so are $(1-\gamma_A)$ and $(1-\gamma_B)$. The ratio of two irrational numbers is generically irrational, so the weak inequalities are unlikely to hold with equality.

For profile $(\mathbf{a}, \mathbf{b})$ to be stochastically stable, the inequality must hold for all singleton sets $Q' = \{A_i\} \subset V^A$. So each individual player $A_i \in V^A$ must have more connections to those in his/her own group than $(1-\gamma_A)/\gamma_A$ times the number of connections to those in $V^B$. Note that for these singleton sets, this is just a restriction on individual behaviour, since if action $a$ is not a best response for player $A_i \in V^A$ to the profile where all other Group $A$ players take action $a$ and all Group $B$ players take action $b$, then profile $(\mathbf{a}, \mathbf{b})$ is not even an equilibrium.

However, holding for all singleton subsets of $V^A$ is insufficient for the inequality to hold for all $Q' \subseteq V^A$ as Figure 3.8 below demonstrates. Agent $A_i \in V^A$ is adjacent to those in the set $V_1^B = \{B_1, B_2\} \subset V^B$. Agent $A_i$ is also adjacent to agents $A_j$ and $A_k$ from $A$. Thus, by taking $Q' = \{A_i\}$, the relevant inequality holds for any value of $\gamma_A$ given that $\gamma_A > 1/2 > 1-\gamma_A$. Similarly, agent $A_j \in V^A$ is adjacent to 2 agents from $V^B$, those in $V_2^B = \{B_3, B_4\} \subset V^B$. Note that $V_1^B \cap V_2^B = \emptyset$. Agent $A_j$ is connected to agents $A_i$, and $A_l$ from Group $A$. So choosing $Q' = \{A_j\}$, inequality (3.16) clearly holds. However, by choosing $Q' = \{A_i, A_j\}$, it is the case that $e(Q', V^B) = 4$, whereas $e(Q', A\backslash Q') = 2$, and as such inequality in (3.16) will only hold for $\gamma_A \in [5/8, 1)$ and not for $\gamma_A \in (1/2, 5/8)$.

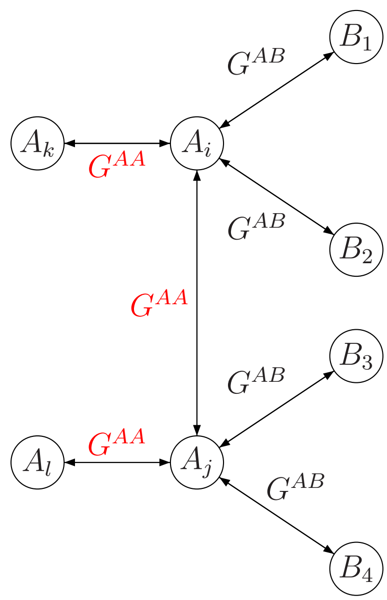


**Figure 3.7**: Failure to coordinate across Group

## 3.9 Conclusion

This paper has examined stochastically stable equilibria to the Language Game of Neary (2010a). While the results are far from conclusive, the take home message should be that network architecture can affect equilibrium selection, and that much of what is driving the risk-dominant predictions of KMR, Ellison (1993), Blume (1993), and Peski (2010) is the homogeneity of the population.

Another interesting feature is how much more sensitive stochastically stable outcomes are to the manner in which players make mistakes. To my knowledge there are no games for which the uniform errors of KMR select different outcomes to the payoff dependent errors of Blume (1993). But with the introduction of multiple populations this immediately comes to the fore. Further exploration of the details of the dynamics seems an interesting avenue for future research.

# Chapter 4

# The Class of Reasonable Equilibria

## Abstract

We propose three new equilibrium refinements for normal form games, that can easily be extended to extensive form analogs. The refinements are strategy-perturbation based, à la perfection, as defined by Selten (1975). An equilibrium is defined as *reasonable* if more costly mistakes are trembled onto with lower probability than less costly ones. However, we do not require that the relative likelihood be of a full order of magnitude, as in properness (Myerson (1978)). It may be, but need not be, the sole requirement is that it be weakly lower. The remaining two refinements are similar in spirit to reasonable equilibria although instead of requiring robustness to one such reasonable tremble, they require robustness to more, and all, such reasonable trembles respectively. We call an equilibrium *more-than-proper* if it is robust to a larger, yet tractable, class of *reasonably* constrained games than properness. We call an equilibrium *truly proper* if it is robust to the set of all reasonably constrained games.

## 4.1 Introduction

The ideas of equilibrium refinements originated with the seminal work of Selten (1975). Selten (1975) observed that in extensive form games the Nash solution concept could often prescribe seemingly *irrational* (non-maximizing) behavior at unreached information sets, and he sought to eliminate such *unreasonable* equilibria. In standard economic models these *unreasonable* equilibria often amounted to players making incredible threats. This produced a huge inconsistency between game theoretical predictions and *reasonable* (or at least observed) economic behaviour, and have caused some concern to game theorists for some time.

Selten (1975) introduced the concept, that assigns a non-empty set of strategy profiles each element of which is known as a *perfect equilibrium point*, by postulating that full rationality is the limiting case of *incomplete rationality*. To quote Okada (1981): "He (Selten) parameterized incomplete rationality by a small positive probability for the breakdown of rationality and defined the perfect equilibrium point to be stable against arbitrarily slight parameterized incomplete rationality".

Selten's goal was to require that players prescribe rational behavior at every information set. Mathematically, his mechanism found an equilibrium point stable against an arbitrary sequence of perturbations in the strategy choice of players. It did this by looking at *nearby* constrained games where all players were required to play each pure strategy with non-zero probability. This ensured players now would (not just could) be called on to act at each information set, hence proposing a non-maximizing action at any information set was irrational.

Note, however, that behavior in these constrained games is fully rational by the standard definition of rationality (maximizing expected payoffs given beliefs). Thus, what Okada (1981) refers to as *incomplete rationality* is really an awareness on the part of each player that his opponents and the player himself are constrained in their allowable choices. These can be viewed as physical constraints placed on the players or (perhaps more plausibly) as *belief constraints* on how a player may make slight mistakes. Perfection required that an equilibrium be robust against an arbitrary set of belief constraints.

The fundamental question that this paper seeks to address is: 'What is the *right* set of constrained games to look at?' Or to put it another way: 'What constrained games are reasonable, and why?' What is unique about our contribution is that we turn this question on its head. Rather we ask: 'What beliefs are unreasonable?' and eliminate these as possible candidates. The remaining set of beliefs are considered possible, and any equilibrium supporting beliefs from this set is referred to as *reasonable*.

It was viewed as undesirable that the process towards full rationality may violate dominance. It seemed odd that players could collectively agree that if strategy $B$ dominates strategy $A$ for player $i$, then player $i$ may tremble onto $A$ with greater probability than $B$. Myerson (1978) noticed this undesirable property of some perfect equilibria, and so proposed a strict refinement of perfection, namely properness. Proper equilibria are perfect equilibria where the added assumption is that there is a particular rational mechanism underlying the perturbations.[1]

Myerson (1978) motivated properness with the idea that players ought to take care as to how they tremble, and as such should tremble onto more costly mistakes with infinitely less probability than onto less costly ones. The main concern is that while players tremble in a reasonable (non-dominance violating) manner, the weighting scheme they assign to strategies is very specific, and in some instances overly restrictive.

The first refinement we present here is a coarsening of properness. We look at constrained games where every player must play a fully mixed strategy, and any pure strategy that is not a best response given beliefs must receive infinitely less weight than all pure strategy best responses. Within the complement of his best response set,[2] a player must give less weight to worse strategies, but need not give infinitely less weight, just less. We also require that payoff equivalent strategies receive equal weight, be they best responses or not. An equilibrium is *reasonable* if it is robust to one such constrained game. Existence is immediate by Jurado and Sánchez (1990), who extend Myerson's properness to also require the equal weighting of payoff equivalent strategies. [3]

[1] We will use "trembles", "perturbations", and "belief constraints" interchangeably throughout.

[2] We will refer to a generic agent as "he" throughout the paper.

[3] Jurado and Sánchez (1990) call their refinement *strong properness*. The difference is mild,

Our next two procedures strictly refine the set of reasonable equilibria. *More-than-proper* equilibria require robustness to the set of reasonably constrained games where every player must play a fully mixed strategy, and any strategy that is not a best response given beliefs must receive infinitely less weight than all pure best responses. Within the complement of the best response set, worse strategies receive a constant $\delta$ times the weight of the next best strategy, and this is required to hold for all $\delta \in (0, 1)$. *Truly proper* equilibria require robustness to all reasonably constrained games. Thus an equilibrium is *truly proper* if it is stable against all sequences of nearby constrained games where the belief constraints do not violate dominance given knowledge of how the opponents are playing. What it means formally for players to have collectively reasonable belief constraints will be made precise in Section 4.5. Truly proper equilibria are a strict refinement of more-than-proper equilibria, but the difference is mild and comes at a cost as more-than-properness is much more tractable, and can even be shown to eliminate some *intuitively bad* proper equilibria. More-than-properness also permits a simple test for selecting between multiple proper equilibria that is discussed in Section 4.7.

The paper is organized as follows: Section 4.2 shows via a simple one player example the differences between each of our 3 new equilibrium refinements, and contrasts them with existing refinements like perfection and properness. Section 4.4 provides a brief recap of normal form equilibrium refinements, highlighting the merits and deficiencies of each. Section 4.5 argues that reasonable trembles are the correct way to think about Selten's so-called *breakdowns in rationality*, and as such properness is a step in the right direction. Section 4.6 argues that the class of allowable reasonable trembles can also be reduced. Section 4.7 formalises our equilibrium notion of *more-than-properness*, while Section 4.8 introduces *true properness*. Section 4.9 presents our main result which categorizes the positioning of our new refinements in the *food chain* of existing refinements. Section 4.10 concludes. Our notation is reasonably standard and is quickly laid out in Section 4.3.

---

so from this point on in the paper whenever we speak of properness, we are referring to strong properness.

## 4.2 A one player Example

In their paper on sequential equilibria, Kreps and Wilson (1982) write: "Some sequential equilibria are supported by beliefs that the analyst can reject because they are supported by beliefs that are implausible. We will not propose any formal criteria for 'plausible beliefs' here." The purpose of this section is to show through a simple example what we propose as *plausible beliefs* that players may hold. This is done by looking at *reasonable* mechanisms by which a rational player could *tremble*.

We consider a set of complete information 1-player normal form games, $\{\Gamma(k, l)\}$, where $k$ and $l$ parameterize the set. $k$ is a real number strictly less than 2, while $l$ is a real number strictly less than $k$.

The story goes like this: Rowena adores Apples ($A$); she quite likes Bananas ($B$); she also thinks Cheese ($C$) is ok, but not as nice as Bananas, while Dates ($D$) are her least favourite food. Suppose Rowena walks into a room with a table on which lies 1 Apple ($A$), 1 banana ($B$), 2 identical pieces of cheese ($C_1, C_2$), and a bag of dates ($D$). Rowena may choose only one item from the table. Her goal is to choose whatever makes her happiest, and she is free to do so. A payoff matrix, $\Gamma(k, l)$, representing this situation is the following:

| | | |
|---|---|---|
| | $A$ | 100 |
| | $B$ | 2 |
| Row | $C_1$ | $k$ |
| | $C_2$ | $k$ |
| | $D$ | $l$ |

First of all notice that while this is technically a game, it is a 1-player game and is more accurately viewed as a decision problem. Second, notice that in the realm of decision problems, this is a trivial one: it is obvious what Rowena should do - she should choose the apple ($A$). Thus the only equilibrium (if it can be called that) is for Rowena to play $A$. Given that $A$ is the only equilibrium and that it is a strict equilibrium,[4] it must therefore be the only robust equilibrium. So our focus here is not on refining the set of equilibria, but rather in demonstrating various *reasonable* mechanisms illustrating how Rowena might make mistakes.

Let $p_i$ be the weight that Rowena gives pure strategy $i \in \{A, B, C_1, C_2, D\}$. In terms of mixed strategies, the equilibrium in this game can be described by the 5-dimensional vector $(p_A, p_B, p_{C_1}, p_{C_2}, p_D) = (1, 0, 0, 0, 0)$. We will analyse how Rowena will make mistakes given the beliefs that she has.[5]

Under perfection, Selten (1975) would allow any sequence of perturbations of $(1, 0, 0, 0, 0)$ that converges to it. In one possible constrained game, Rowena could play a totally mixed strategy described by $(1-7\varepsilon, \varepsilon, 2\varepsilon, 3\varepsilon, \varepsilon)$. We think this is odd for at least two reasons. First, given that $k < 2$ it seems odd that Rowena would put twice as much weight on $C_1$ than $B$, and three times as much weight on $C_2$ that $B$. Second, given that $C_1$ and $C_2$ represent duplications of the same thing, it is strange that Rowena would place more weight on $C_2$.

Under properness, Myerson (1978) argued that worse strategies, equivalently more costly mistakes, should receive infinitely less weight in the limit than better ones. Thus Myerson might insist on a mixed strategy of the form $(1, \varepsilon, \varepsilon^2, \varepsilon^2, \varepsilon^3)$.[6] This seems plausible if Cheese ($C$) is really unpleasant to Rowena while bananas ($B$) are nice, just not nearly as nice as Apples ($A$). But suppose that while Rowena strictly prefers Bananas to Cheese, she does so only slightly. Then requiring a full order of magnitude weighting difference seems excessive.

Under reasonableness, we argue that while $C$ is clearly worse than $B$ the relative weighting of $B$ to $C$ (and also $C$ relative to $D$) could very well be context specific, and while $C$ should certainly not receive more weight than $B$, insisting that it get infinitely less is in many cases overly restrictive (the case where $k$ is strictly less than but very close to 2 is a clear example).[7] Reasonableness also requires that

[4] A strict equilibrium is one in which each player has a unique best response against the profile. Clearly it must be in pure strategies.

[5] Note now why we discuss this in a decision problem sense. Rowena can have only one belief - that she is allowed play the game (choose an item) with certainty.

[6] In fact Myerson's properness would allow $(1, \varepsilon, \varepsilon^2, \varepsilon^2/2, \varepsilon^3/2)$, but this is ruled out under *strong properness* (Jurado and Sánchez (1990)).

[7] Objections to this by appealing to the cardinal nature of von-Neumann Morgenstern payoffs

duplicate strategies receive equal weight, and that the sum of weights on duplicate strategies does not exceed the weight assigned to any of the next best strategies. We have that the set of $\varepsilon$-reasonable equilibria are $\{(p_A, p_B, p_{C_1}, p_{C_2}, p_D) \,|\, p_A + p_B + p_{C_1} + p_{C_2} + p_D = 1\,;\, p_A, p_B, p_{C_1}, p_{C_2}, p_D > 0\,;\, p_A < 1\,;\, p_B \leq \varepsilon p_A\,;\, p_D \leq p_{C_1} = p_{C_2} \leq p_B/2\}$. Clearly as $\varepsilon$ goes to zero, any sequence in this set converges to $(1, 0, 0, 0, 0)$.

Our next refinement, more-than-properness, requires robustness to a larger class of constraints than properness. More-than-properness assumes that the modeler is aware of how the players would rank their strategies given their beliefs, but is unsure as to how much worse certain strategies are than others. The modeler also assumes that in the complement of the best response set, that worse strategies get a constant fraction of the weight of next best strategies. The modeler assumes this constant weighting scheme, but since they are unsure as to which type of constant relative trembler the players are, requires robustness to all such constant relative weighting schemes. For our example, the set of constrained games is the following: $\{(p_A, p_B, p_{C_1}, p_{C_2}, p_D) \,|\, p_A + p_B + p_{C_1} + p_{C_2} + p_D = 1\,;\, p_A, p_B, p_{C_1}, p_{C_2}, p_D > 0\,;\, p_A < 1\,;\, p_B \leq \varepsilon p_A\,;\, p_{C_1} = p_{C_2} = \delta p_B/2\,,\, p_D = \delta p_{C_1} = \delta p_{C_2}\}$. Here, $\delta$ captures the idea of constant relative trembling. As $\varepsilon$ goes to zero, for a given value of $\delta$ we have a $\delta$-more-than-proper equilibrium. An equilibrium is said to be more-than-proper if it is $\delta$-more-than-proper for all values of $\delta \in (0, 1)$.

Finally, we also propose the concept of true-properness. This is done to deflect two possible objections to more-than-properness. First, more-than-properness has the potentially undesirable feature that while the modeler is unsure as to what type (what value of $\delta$) of trembler each player is, he assumes that all players are simultaneously of the same type. But surely it is possible that different players may *tremble* in different ways. Second, the constant relative trembling is attractive due to its tractability (as it can be parameterized by only one parameter), but seems excessive. Letting $\underline{\delta} = (\delta', \delta'')$ be a vector in $(0, 1) \times (0, 1)$, the set of $\underline{\delta}$-truly proper equilibria is the following: $\{(p_A, p_B, p_{C_1}, p_{C_2}, p_D) \,|\, p_A + p_B + p_{C_1} + p_{C_2} + p_D = 1\,;\, p_A, p_B, p_{C_1}, p_{C_2}, p_D > 0\,;\, p_A < 1\,;\, p_B \leq \varepsilon p_A\,;\, p_{C_1} = p_{C_2} = \delta' p_B/2\,,\, p_D =$

which can make any two close payoffs arbitrarily distant from each other can equally be used as a critique of properness too. That is, if two payoffs can be made arbitrarily far apart, they can also be made arbitrarily close by the reverse transformation.

$\delta'' p_{C_1} = \delta'' p_{C_2}\}$. The equilibrium is truly proper if it is $\underline{\delta}$-truly proper for all pairs $(\delta', \delta'') \in (0,1) \times (0,1)$.

The one player decision problem above does not capture how much stronger the notion of true properness is as compared to more-than-propernes. For multi-player games, true properness says that given a particular type of trembler for player $i$, the constrained equilibrium must be robust for all types of relative tremblers of the other players. And this statement must hold no matter what type of relative trembler player $i$ is.

## 4.3 Notation

A finite $n$-person normal form game is a $2n$-tuple $\Gamma = (S_1, \ldots, S_N, u_1, \ldots, u_n)$, where $S_i$ is a finite nonempty set of pure strategies for player $i$, and $u_i$ is player $i$'s payoff function defined as a mapping $u_i : \prod_{j=1}^{n} S_j \longrightarrow \mathbb{R}$. These hold for $i \in \{1, \ldots, n\}$.

$\Sigma_i$ denotes the set of all probability distributions over $S_i$, and we call this set player $i$'s mixed strategies space. We define $\mathbf{S} := \prod_{j=1}^{n} S_j$ and $\boldsymbol{\Sigma} := \prod_{j=1}^{n} \Sigma_j$, with typical elements $\mathbf{s} = (s_1, \ldots, s_N)$, and $\boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_n)$. It is straightforward to extend $u_i$ to $\boldsymbol{\Sigma}$.

From player $i$'s perspective a pure strategy profile $\mathbf{s} \in \mathbf{S}$ can be viewed as $\mathbf{s} = (\mathbf{s} \backslash s_i)$, while a mixed strategy profile $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ can be viewed as $\boldsymbol{\sigma} = (\boldsymbol{\sigma} \backslash \sigma_i)$. Thus the strategy profile $(\mathbf{s} \backslash \hat{s}_i)$ will refer to $(s_1, \ldots, s_{i-1}, \hat{s}_i, s_{i+1}, \ldots, s_n)$, i.e. the profile $\mathbf{s}$, where $\hat{s}_i$ replaces $s_i$. Similarly for $(\boldsymbol{\sigma} \backslash \hat{\sigma}_i)$, and $(\boldsymbol{\sigma} \backslash \hat{s}_i)$.

We will denote the set of pure best responses for player $i$ to strategy profile $\boldsymbol{\sigma}$ as $\text{BR}_i(\boldsymbol{\sigma})$, where $\text{BR}_i : \boldsymbol{\Sigma} \twoheadrightarrow S_i$. We denote the set of all pure best replies to strategy profile $\boldsymbol{\sigma}$ by $\text{BR}(\boldsymbol{\sigma}) \equiv \prod_{i=1}^{n} \text{BR}_i(\boldsymbol{\sigma})$.

A constrained game (to be made precise below) will be denoted $\Gamma^{\varepsilon}$, where the superscript $\varepsilon$ (which may be a vector) fully parameterizes the constraints. Finally, we will denote the set of equilibria of a game $\Gamma$ by $E(\Gamma)$.

## 4.4 Recap of Definitions

The notion of a constrained game is fundamental to the strategy perturbation based refinements literature. The constraints require that each player places strictly positive weight on each pure strategy, thus guaranteeing that in an extensive form game every information set is reached.[8] This ingenious trick of Selten's was intended to lead to rational behaviour being prescribed at every information set.[9] He viewed constraints as corresponding to mistake probabilities that all players $j \neq i$ hold about player $i$'s actions.[10]

**Definition 22.** Let $\Gamma = (S_1, \ldots, S_N, u_1, \ldots, u_n)$ be an $n$-person finite normal form game. For each $i$, let $\varepsilon_i$ and $\Sigma_i^{\varepsilon_i}$ be defined by:

$$\varepsilon_i \in \mathbb{R}^{S_i} \text{ with } \varepsilon_i(s_i) > 0 \text{ for all } s_i \in S_i, \text{ and } \sum_{s_i \in S_i} \varepsilon_i(s_i) < 1, \tag{4.1}$$

$$\Sigma_i^{\varepsilon_i} := \Big\{ \sigma_i \in \Sigma_i : \sigma_i(s_i) \geq \varepsilon_i(s_i), \;\; \forall\, s_i \in S_i \Big\} \tag{4.2}$$

Furthermore, let $\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_n\}$ and $\boldsymbol{\Sigma}^{\varepsilon} := \prod_{i=1}^{n} \Sigma_i^{\varepsilon_i}$. The *constrained game* $\Gamma^{\varepsilon}$ is thus the infinite normal form game $(\Sigma_1(\varepsilon_1), \ldots, \Sigma_n(\varepsilon_n), u_1, \ldots, u_n)$.

Loosely put, there are three "types" of existing strategy perturbation point-valued solution concepts for normal form games, respectively due to Selten (1975), Myerson (1978), and Okada (1981):

---

[8]Our definitions will be for normal form games throughout. This is to avoid having to specify purely extensive form notions like 'sequential best responses', etc. One may think of our analysis in the standard way of each player optimizing *ex-ante*. Each definition can of course be extended to general extensive form games.

[9]By the standard definition of rational behavior, that of maximizing expected payoffs given (possibly constrained) beliefs, this does lead to rational behavior. However it can lead to what we later term unreasonable behavior.

[10]A further assumption is that player $i$ is also aware of the mistakes he himself will make. See Van Damme (1984) for the definition of a quasi-perfect equilibrium point, i.e. a strategy profile that is the limit of completely mixed strategy profiles in which each player is aware of the mistake possibilities of other, but does not take into account that he himself will make a mistakes in the future.

**(1) Perfection:** A strategy profile $\boldsymbol{\sigma}$ is called *perfect*, if there exists at least one sequence of constrained games $\{\Gamma^\varepsilon\}$, and while $\varepsilon \downarrow 0$, we have $\boldsymbol{\sigma}^\varepsilon \in E(\Gamma^\varepsilon)$, and $\boldsymbol{\sigma}^\varepsilon \to \boldsymbol{\sigma}$.

**(2) Properness:** A strategy profile $\boldsymbol{\sigma}$ is called *proper*, if there exists at least one "nice and intuitively reasonable" sequence of constrained games $\{\Gamma^\varepsilon\}$, and while $\varepsilon \downarrow 0$, we have $\boldsymbol{\sigma}^\varepsilon \in E(\Gamma^\varepsilon)$, and $\boldsymbol{\sigma}^\varepsilon \to \boldsymbol{\sigma}$.

**(3) Strict Perfection:** A strategy profile $\boldsymbol{\sigma}$ is called *strictly perfect*, only if for all constrained games $\{\Gamma^\varepsilon\}$, while $\varepsilon \downarrow 0$, we have $\boldsymbol{\sigma}^\varepsilon \in E(\Gamma^\varepsilon)$, and $\boldsymbol{\sigma}^\varepsilon \to \boldsymbol{\sigma}$.

The major contribution of Selten (1975) was to show that no matter what the constraints, an equilibrium in the constrained game "passes to the limit", and that this limit is an equilibrium of the original game. The real power of perfection is that in a constrained game a weakly dominated strategy becomes strictly so, due to every information set being reached with strictly positive probability, and as a rational (payoff maximizing) player will assign as little weight as is permitted to weakly dominated strategies. In passing to the limit, a perfect equilibrium therefore places zero weight on weakly dominated strategies. Kohlberg and Mertens (1986) argue forcefully that players should assign zero weight to weakly dominated strategies. So perfection is certainly a step in the right direction.

The definition of perfection we provide below is due to Myerson (1978). This equivalent way of formalizing a perfect equilibrium introduces the notion of an $\varepsilon$-perfect equilibrium.[11] Thus we can consider perfect equilibria from the perspective of convergence of $\varepsilon$-perfect equilibria.[12]

**Definition 23.** A strategy profile $\boldsymbol{\sigma}^\varepsilon \in \boldsymbol{\Sigma}$ is said to be an $\varepsilon$-perfect equilibrium of $\Gamma$, if $\boldsymbol{\sigma}^\varepsilon$ is completely mixed, and satisfies the following:

[11] Intuitively, this just says that if a response to a strategy profile is not a best response for player $i$, then it must get weight of no more than $\varepsilon$. It can be shown that an equilibrium is perfect if and only if it is a limit point of $\varepsilon$-perfect equilibria - see Theorem 8.4 (page 352) in Fudenberg and Tirole (1991).

[12] In an abuse of notation, $\varepsilon$ may refer to a real number strictly between 0 and 1, or may refer to a vector of constraints. It should always be clear from the context.

if $\quad u_i(\boldsymbol{\sigma}^\varepsilon \backslash s_i') < u_i(\boldsymbol{\sigma}^\varepsilon \backslash s_i'')$,

then $\quad \sigma_i^\varepsilon(s_i') \leq \varepsilon$, for all $i$, for all $s_i', s_i''$

A perfect equilibrium is the limit point of a sequence $\{\boldsymbol{\sigma}^\varepsilon\}_{\varepsilon\downarrow 0}$, where $\boldsymbol{\sigma}^\varepsilon$ is an $\varepsilon$-perfect equilibrium of $\Gamma$, for all $\varepsilon$.

The weakness of perfection is that for an equilibrium to be stable that it be robust to some slight strategy perturbation. But the perturbation can be chosen arbitrarily - its sole purpose is that it "works".

Myerson (1978) introduced the notion of a proper equilibrium point which implicitly assumes that players will try harder to avoid more costly mistakes. Properness has been interpreted as there being some rational mechanism governing the mistake probabilities. Strategies that are more costly are assigned a probability of smaller order than less costly ones. Formally, under perfection, players cannot act irrationally by the definition of rational behavior (prescribe non-maximizing behaviour), but can hold and respond to what are deemed *unreasonable* beliefs over the mistake probabilities of others. Under properness, mistake probabilities of players are constrained in a reasonable, but very specific way. For completeness the formal definition is given below.

**Definition 24.** Let $\Gamma$ be an $n$-person normal form game. Let $\varepsilon > 0$ and $\boldsymbol{\sigma}^\varepsilon \in \boldsymbol{\Sigma}$. Myerson (1978) calls $\boldsymbol{\sigma}^\varepsilon$ an *$\varepsilon$-proper equilibrium* of $\Gamma$ if $\boldsymbol{\sigma}^\varepsilon$ is completely mixed and satisfies:

if $\quad u_i(\boldsymbol{\sigma}^\varepsilon \backslash s_i') < u_i(\boldsymbol{\sigma}^\varepsilon \backslash s_i'')$,

then $\quad \sigma_i^\varepsilon(s_i') \leq \varepsilon \sigma_i^\varepsilon(s_i'')$, for all $i$, for all $s_i', s_i''$.

A proper equilibrium is the limit point of a sequence $\{\boldsymbol{\sigma}^\varepsilon\}_{\varepsilon\downarrow 0}$, where $\boldsymbol{\sigma}^\varepsilon$ is an $\varepsilon$-proper equilibrium of $\Gamma$, for all $\varepsilon$.

An $\epsilon$-proper equilibrium always exists in the constrained game. Since each constrained game possesses at least one equilibrium, and this equilibrium is an element of a compact set, the limit of the constrained equilibria (possibly passing to subsequences) exists. That this limit is an equilibrium of the original game is trivial, by showing that any proper equilibrium is perfect. Thus a proper equilibrium exists for all normal form games.

Okada (1981) introduced the notion of a *strictly perfect* equilibrium point. Intuitively this just generalises perfection, i.e. where perfection requires stability to *any one* sequence of nearby constrained games, strict perfection requires stability to *all* sequences of nearby constrained games.

**Definition 25.** Let $\Gamma$ be an $n$-person normal form game. Let $\hat{\eta} \in \mathbb{R}^{\mathbf{S}}_{++}$, let $O_{\hat{\eta}} := \left\{\eta \in \mathbb{R}^{\mathbf{S}}_{++} \mid \eta < \hat{\eta}\right\}$.[13] Call $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ a *strictly perfect equilibrium* of $\Gamma$ if there exists some $\hat{\eta} \in \mathbb{R}^{\mathbf{S}}_{++}$ and for each $\eta \in O_{\hat{\eta}}$ some $\boldsymbol{\sigma}^{\eta} \in E(\Gamma^{\eta})$ such that $\lim_{\eta \downarrow 0} \boldsymbol{\sigma}^{\eta} = \boldsymbol{\sigma}$.

Van Damme (1996) introduced the following refinement of strict perfection where a mild regularity condition is placed on the strategy perturbations. This condition ensures that strategies change smoothly as beliefs change. He refers to such an equilibrium as *strictly proper*.

**Definition 26.** An equilibrium $\boldsymbol{\sigma}$ of $\Gamma$ is strictly proper if $\boldsymbol{\sigma}$ is strictly perfect <u>and</u> the mapping $\eta \mapsto \boldsymbol{\sigma}(\eta)$ from $O_{\hat{\eta}}$ to $\boldsymbol{\Sigma}$ is continuous.

In the next Section we will argue that if any behavioral interpretation is attached to the notion of trembling, then surely players collectively believing that another player will violate dominance in his mistakes is *unreasonable*. So only beliefs in the complement of this unreasonable set are plausible. Equilibria supporting supported by such beliefs will be called *reasonable*.

## 4.5 Reasonable Equilibria

That equilibria should be robust to trembles is widely accepted. Interpreting trembles as uncertainty in beliefs rather than a mere technical trick is more controversial.[14] However, if you require robustness to trembles <u>and</u> take seriously the behavioral notion of trembles, then *reasonableness*, as defined below, is surely a necessary condition for an equilibrium to be self enforcing.

[13] Given two equal dimensional vectors $\mathbf{x}$ and $\mathbf{y}$, write $\mathbf{x} < \mathbf{y}$ if $x_i < y_i$ for all $i$.

[14] Harsanyi and Selten (1988) argued that *some* behavioral interpretation should be attached to the notion of trembles, while Kohlberg and Mertens (1986) argue that such an interpretation should not be made.

The common theme to all our refinements is that players do not violate dominance given their beliefs. We insist that players may make mistakes and are aware that other do too. Note that no matter what beliefs a player holds, it is always possible to rank his pure strategies in terms of the payoff each would yield. We insist that best responses to a particular belief receive *most of the* weight, while non-best responses receive *little* weight. Inferior non-best responses receive less weight than better non-best responses.

It seems highly believable that players will make *less costly* mistakes. This is easily understood and quite intuitive as the following suggests: it seems quite odd to suppose that if pure strategy $A$ and $B$ are both strictly dominated relative to a proposed equilibrium profile, while strategy $B$ strictly dominates $A$, that strategy $A$ should be assigned more weight than $B$ in any constrained game. If not the payoffs are surely not representative of the player's preferences.

However, it also seems overly restrictive to insist that the weight put on a more costly mistake always be a full order of magnitude lower than that on a less costly one as in properness. Since if a strategy is only marginally less costly then the trembling player is exerting a huge amount of energy to control the less costly mistake despite it being only slightly so. Furthermore, there may be external factors not present in the normal form, such that despite a particular player having a strategy appear much more costly than another to the modeler, perhaps the player has an unobservable fondness for that particular strategy. Provided the player is rational this more costly error will certainly receive less weight, but need not receive a full order of magnitude less weight as in the properness notion of Myerson (1978).

The deficiencies of perfection are highlighted in the game 4.5 below. The equilibrium $\{L_1, r_2\}$ is perfect: Provided player 2 believes that $R_1$ is more than twice as likely a 'mistake' than $M_1$, then he is justified in playing $r_2$. But this is absurd as $M_1$ strictly dominates $R_1$ for player 1, and as such player 2 is holding an unreasonable belief if he thinks that player 1 may play $R_1$ with higher probability than $M_1$. Surely if player 2 is called upon to act he will believe it more likely that player 1 played $M_1$ rather than $R_1$. Thus we believe the set of constrained games that require player 1 to put strictly more weight on $R_1$ than $M_1$ to be very unreasonable.

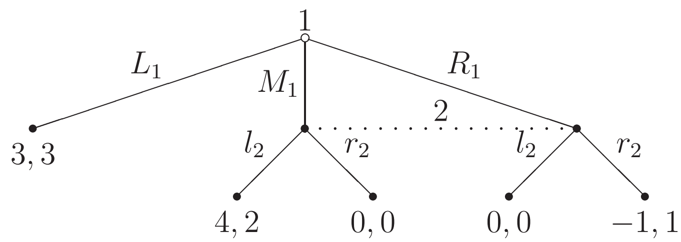


**Figure 4.1**: Unreasonable belief constraints

Thus, taking the view that trembles are restrictions on players beliefs, it seems overly restrictive and intuitively quite odd to require that complete rationality be viewed as the limiting case of any kind of unreasonable beliefs, as these are surely a component of unreasonable behavior. We believe that incomplete rationality, of which complete rationality is the limit, would better be modeled as having not-quite-certain but still reasonable belief constraints. It seems far more natural to model the complete rationality of an individual as the limiting case of nearby games in which players can tremble in any manner they wish as long as it is reasonable. We now formally define what it means for an equilibrium to satisfy this limit of reasonable beliefs. Naturally such equilibria are referred to as *reasonable.*

First we will need the definition of an equivalence class.[15]

**Definition 27.** Given a strategy set $S_i$ for any player $i$. Define the *equivalence class* of strategy $s_i \in S_i$ as the subset of all elements of $S_i$ that are payoff equivalent to $s_i$:

$$[s_i] := \{s_i' \in S_i \,|\, u_i(\boldsymbol{\sigma} \backslash s_i') = u_i(\boldsymbol{\sigma} \backslash s_i), \forall\, \boldsymbol{\sigma} \in \boldsymbol{\Sigma}\}$$

If $s_i'$ and $s_i''$ are both elements of the same equivalence class, we will write $s_i' \sim s_i''$. The number of elements in $[s_i]$ will be denoted by $|[s_i]|$.

[15] An equivalence relation is a binary relation (often denoted by $\sim$) on a set $X$. The following should be read with the appropriate quantifiers "for all $x, y, z$," etc. That is, a relation that is reflexive ($x \sim x$), symmetric ($x \sim y$ implies $y \sim x$), and transitive ($x \sim y$ and $y \sim z$ imply $x \sim z$). Given an equivalence relation $\sim$ on a set $X$, we define the equivalence class $[x]$ of $x$ by $[x] := \{y : y \sim x\}$.

**Definition 28.** Let $\Gamma$ be an $n$-person normal form game. Let $\varepsilon > 0$ and $\boldsymbol{\sigma}^\varepsilon \in \boldsymbol{\Sigma}$. We call $\boldsymbol{\sigma}^\varepsilon$ an *$\varepsilon$-reasonable equilibrium* of $\Gamma$ if $\boldsymbol{\sigma}^\varepsilon$ is completely mixed and satisfies:

1. if $\quad s_i \in \text{BR}(\boldsymbol{\sigma}^\varepsilon)$,
   then $\quad \sigma_i^\varepsilon(s_i') \leq \varepsilon \sigma_i^\varepsilon(s_i)$, for all $i$, for all $s_i' \notin \text{BR}(\boldsymbol{\sigma}^\varepsilon)$.

2. if $\quad u_i(\boldsymbol{\sigma}^\varepsilon \backslash s_i') < u_i(\boldsymbol{\sigma}^\varepsilon \backslash s_i'')$,
   then $\quad \sigma_i^\varepsilon(s_i') \leq \frac{1}{|[s_i']|} \sigma_i^\varepsilon(s_i'')$, for all $i$, for all $s_i', s_i'' \notin \text{BR}(\boldsymbol{\sigma}^\varepsilon)$.

3. if $\quad s_i' \sim s_i''$,
   then $\quad \sigma_i^\varepsilon(s_i') = \sigma_i^\varepsilon(s_i'')$, for all $i$.

A reasonable equilibrium is the limit point of a sequence $\{\boldsymbol{\sigma}^\varepsilon\}_{\varepsilon \downarrow 0}$, where $\boldsymbol{\sigma}^\varepsilon$ is an $\varepsilon$-reasonable equilibrium of $\Gamma$, for all $\varepsilon$.

Some remarks are in order:

*Remark* 1. Condition 1 in the definition follows Myerson in that it requires pure strategies that are not best responses to the group profile receive at most $\varepsilon$ times the weight of those pure strategies that are best responses. Condition 2 only requires that within the complement of their best response set, that no player violates dominance with their trembles. Condition 3 just ensures that payoff equivalent strategies are assigned equal weight.

*Remark* 2. The introduction of an equivalence class for each strategy is necessary as without it, the solution concept is sensitive to the duplication of strategies. Without the scaling in condition 2, one could continue to add inferior strategies and their combined weight could eventually exceed that of a "*better*" strategy.

*Remark* 3. The reader may object to a strictly worse strategy being allowed as much weight as a better one. This is easily handled by a slight tweaking definition 28, where instead of defining an $\varepsilon$-reasonable equilibrium, one instead defines an $(\varepsilon\delta)$-reasonable equilibrium, where $\delta$ is strictly less than (but close to) 1, and is inserted on the right hand side of the weak inequality in condition 2.

Thus, in Game 4.5 above, the perfect equilibrium $\{L_1, r_2\}$ is not self-enforcing as it is not reasonable. Since player 1 ranks his strategies given his beliefs and assigns them weight in an order preserving manner, there is no belief that player 1 can ever hold that will lead player 2 to believe that $R_1$ is a more likely 'mistake' than $M_1$. Thus player 2 must play $l_2$.

Reasonableness restricts the set of perturbations that are allowed. We will denote the set of reasonable equilibria of a game $\Gamma$ by $Rea(\Gamma)$.

**Lemma 11.** *$Rea(\Gamma)$ is a nonempty closed subset of the set of perfect equilibria. Every proper equilibrium is reasonable.*

*Proof.* That every reasonable equilibrium is perfect is trivial by definition 23.

Existence is trivial since it is clear from definition 24 that every proper equilibrium is reasonable, and proper equilibria always exist.

To show that $Rea(\Gamma)$ is closed we proceed as follows. Fix $\varepsilon > 0$. Let $R_\varepsilon$ be the set of $\gamma$-reasonable equilibria for $\gamma \in (0, \epsilon)$. Let $\overline{R}_\varepsilon$ denote the closure of $R_\varepsilon$. Clearly $Rea(\Gamma) = \cap_{\varepsilon>0}\overline{R}_\varepsilon$. Since $\overline{R}_\varepsilon$ is closed and the intersection of any arbitrary collection of closed sets is closed we are done. □

Consider the following game $\Gamma(x)$, similar to one in Van Damme (1996):

| | | $P_2$: L | C | R |
|---|---|---|---|---|
| $P_1$ | U | $2,3$ | $1,2$ | $0,1$ |
| | M | $2,3$ | $0,2$ | $x,1$ |

It is clear that Player 2 has a strictly dominant strategy that he must play, thus any equilibrium requires that he play $L$. In fact, player 2 has a strict preference ordering on his set of pure strategies. Let $p$ denote the probability that player 1 plays $U$. Note that every equilibrium in the connected set of equilibria $\{\{p, L\} : p \in [0, 1]\}$ is reasonable.[16] It is also not difficult to see that the unique proper equilibrium to this game is $\{U, L\}$. However, $\{U, L\}$ is only proper because player 2 strictly prefers $C$ to $R$ and as such is required to play $R$ with a full order of

[16] Let $X$ be a topological space, and $A$ a subset of $X$. $A$ is said to be *connected* if it cannot be represented by the union of two or more disjoint nonempty subsets each of which is open in the relative topology induced on $A$.

magnitude less than $C$. Player 1's best response to this constrained optimization is to play $U$ since given the order of magnitude of player 2's tremble, $U$ is better than $M$ (i.e. $1\epsilon > x\epsilon^2$ for all $\epsilon$ sufficiently close to zero no matter what the value of $x$).

## 4.6 But is Reasonableness enough?

The following process shows the need for an extension to properness (and hence reasonableness by Theorem 11).[17] Consider the following 2-player normal form game:

| $P_1$ \ $P_2$ | L | C |
|---|---|---|
| U | 4, 4 | 2, 2 |
| M | 2, 2 | 2, 2 |

**Figure 4.2**: A nongeneric game with 2 (pure strategy) equilibria.

This is a non-generic game with only 2 equilibria. Both equilibria, $\{U, L\}$ and $\{M, C\}$, are in pure strategies. As discussed before, since perfection is a rather weak requirement, if an equilibrium is not perfect then it is very unstable. However, as is easily shown, the only perfect equilibria in this game is $\{U, L\}$. In Figure 4.6 below, we transform the game by adding a weakly dominated strategy to each player.

| $P_1$ \ $P_2$ | L | C | R |
|---|---|---|---|
| U | 4, 4 | 2, 2 | 0, 0 |
| M | 2, 2 | 2, 2 | 2, 2 |
| D | 0, 0 | 2, 2 | 2, 2 |

**Figure 4.3**: Game above altered by adding weakly dominated strategies.

[17] Our development in this section follows closely that of Myerson (1978).

Thus we have added a weakly dominated strategy to each player and it can now be seen that both $\{U, L\}$ and $\{M, C\}$ are perfect. In fact, the equilibrium $\{M, C\}$ is reasonable, and even proper.[18] Thus even properness has selected the unappealing equilibrium $\{M, C\}$.

## 4.7 More-Than-Proper Equilibria

We believe that the right set of constrained games to focus on is the set of all reasonably constrained games. More-than-properness is arrived at by looking at a particular subset of these reasonably constrained games. More-than-proper equilibria are the limit of a sequence of $(\varepsilon\delta)$-more-than-proper equilibria. Before defining this, we will need one additional piece of notation.

The following definition ranks a player's strategies relative to a given strategy profile.

**Definition 29.** The number of pure strategies of player $i$ strictly preferred to $\hat{s}_i$ relative to the profile $\boldsymbol{\sigma}$ is

$$c_i(\boldsymbol{\sigma}, \hat{s}_i) := \#\Big\{ s_i \in S_i \,|\, u_i(\boldsymbol{\sigma} \backslash s_i) > u_i(\boldsymbol{\sigma} \backslash \hat{s}_i) \Big\} \tag{4.3}$$

Clearly if $\hat{s}_i \in \text{BR}_i(\boldsymbol{\sigma})$, then $c_i(\boldsymbol{\sigma}, \hat{s}_i) = 0$, and if $s'_i \sim s''_i$, then $c_i(\boldsymbol{\sigma}, s'_i) = c_i(\boldsymbol{\sigma}, s''_i)$.

**Definition 30.** Let $\Gamma = (S_1, \ldots, S_N, u_1, \ldots, u_n)$ be an $n$-person normal form game. Let $\varepsilon > 0$, and $\delta \in (0, 1)$, and $\boldsymbol{\sigma}^{\varepsilon\delta} \in \boldsymbol{\Sigma}$. We call $\boldsymbol{\sigma}^{\varepsilon\delta}$ an $(\varepsilon\delta)$-*more-than-proper equilibrium* of $\Gamma$ if it is completely mixed and satisfies:

1. if $\quad s_i \in \text{BR}_i(\boldsymbol{\sigma}^{\varepsilon\delta})$,
   then $\quad \sigma_i^{\varepsilon\delta}(s'_i) \leq \varepsilon \sigma_i^{\varepsilon\delta}(s_i)$, for all $i$, for all $s'_i \notin \text{BR}_i(\boldsymbol{\sigma}^{\varepsilon\delta})$.

2. if $\quad c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s''_i) < c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s'_i)$, and $\nexists\, \tilde{s}_i \in S_i$ such that
   $c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s''_i) < c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, \tilde{s}_i) < c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s'_i)$,
   then $\quad \sigma_i^{\varepsilon\delta}(s'_i) = \frac{1}{|[s'_i]|} \delta \sigma_i^{\varepsilon\delta}(s''_i)$, for all $i$, for all $s'_i, s''_i \notin \text{BR}_i(\boldsymbol{\sigma}^{\varepsilon\delta})$

[18] $\big(\sigma_1(U), \sigma_1(M), \sigma_1(D)\big) = \big(\varepsilon^2, 1, \varepsilon\big)$ does it for player 2 to be willing to randomize in a similar manner. The game is symmetric so only needs to be checked for one player.

3. if $s_i', s_i'' \notin \text{BR}_i(\boldsymbol{\sigma}^{\varepsilon\delta})$, and $s_i', s_i''$ are payoff equivalent
   then $\sigma_i^{\varepsilon\delta}(s_i') = \sigma_i^{\varepsilon\delta}(s_i'')$, for all $i$.

$\boldsymbol{\sigma}^{\delta} \in \boldsymbol{\Sigma}$ is a $\delta$-more-than-proper equilium of $\Gamma$, if $\boldsymbol{\sigma}^{\delta}$ is a limit point of a sequence $\left\{\boldsymbol{\sigma}^{\varepsilon\delta}\right\}_{\varepsilon\downarrow 0}$ where $\boldsymbol{\sigma}^{\varepsilon\delta}$ is an $\varepsilon\delta$-proper equilibrium of $\Gamma$ for all $\varepsilon$, for some $\delta \in (0,1)$.

Finally we say that $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ is a *more-than-proper equilibrium* of $\Gamma$, if $\boldsymbol{\sigma}$ is a $\delta$-proper equilibrium of $\Gamma$ for all $\delta \in (0,1)$.

Again, some remarks are in order:

*Remark* 4. The equality in the second condition is what gives more-than-properness its bite. To see the difference between more-than-properness and properness consider again game 4.6. Clearly $\{U, L\}$ is proper but it is not as obvious that $\{M, C\}$ can be made proper with the appropriately defined tremble given before. However, only $\{U, L\}$ is more-than-proper. $\{M, C\}$ is $\delta$-proper for values of $\delta \in (0, 1/2]$, but fails to be $\delta$-proper for values of $\delta > 1/2$. Thus our more-than-proper equilibrium concept has removed the less desirable equilibria $\{M, C\}$ from Game 4.6.

*Remark* 5. More-than-proper equilibria do not exist for all games. Refer back to Game 4.5. Note that if we permit player 2 to be an $(\epsilon\delta)$-trembler, then this preference of player 1 for $U$ over $M$ need not always hold. For values of $x > 1$ we run into trouble for $\delta > 1/x$, since in that case it need not be true that $1\epsilon > x\epsilon\delta$. For small values of $\delta$ the $(\epsilon\delta)$-more-than-proper equilibrium coincides with the $\epsilon$-proper equilibrium $\{U, L\}$, but for large values of $\delta$ the $(\epsilon\delta)$-more-than-proper equilibrium flips to $\{M, L\}$. Since more-than-properness requires intersecting over all $\delta \in (0,1)$ we have that the only proper equilibrium of the game $\{U, L\}$ is not more-than-proper. But while more-than-proper equilibria do not exist for all games, the concept could be extended from a point-valued solution to set-valued one. We leave this to future research.

*Remark* 6. While more-than-properness fails to exist for all games, it can in some instances provide a method for selecting amongst multiple proper equilibria.[19] An *index of stability* can be created for how robust a proper equilibrium is. For

[19] I thank Chulyoung Kim for suggesting this approach.

example, in Game 4.6, the proper equilibrium $\{M, C\}$ has only a stability index of $\delta = 1/2$ and thus is not particularly stable, whereas the proper equilibrium $\{U, L\}$ has a stability index of 1. This is as our intuition suggests.

*Remark* 7. It is not difficult to see that as in the definition of an $\varepsilon$-reasonable equilibrium, the use of equivalence classes is again necessary to maintain robustness to duplication of strategies.

*Remark* 8. Phelps and Pollak (1968) rewrote the way economists thought about discounting. The standard models assumed that lifetime utility, $U$, was separable over periods and that the following days felicity (intertemporal utility), $u$, was always a constant discount factor less valuable than the preceding day; i.e. $U(c_0, \ldots, c_T) = \sum_{t=0}^{T} \beta^t u(c_t)$, where $\beta \in [0, 1]$ is the discount factor. This constant discounting scheme is similar to the weighting scheme used by Myerson in properness, where next worse strategies are assigned a constant weight of the next best strategy.. But Phelps and Pollak (1968) conjectured that possibly the immediate following period is *relatively* more important than future periods; i.e. that $U(c_0, c_1, \ldots, c_T) = u(c_0) + \beta \sum_{t=1}^{T} \delta^{t-1} u(c_t)$, where both $\beta, \delta \in [0, 1]$. More-than-properness is an extension of properness somewhat analogous to how Phelps and Pollak (1968) extended the standard exponential discounting model, where the solution must be robust to the entire family of possible discounting mechanisms.

We denote the set of proper equilibria of a game $\Gamma$ by $Prop(\Gamma)$ and the set of more-than-proper equilibria by $MTP(\Gamma)$.

**Lemma 12.** *The set of more-than-proper equilibria, MTP*$(\Gamma)$*, is a closed subset of Prop*$(\Gamma)$*. Though possibly empty.*

*Proof.* Let $\boldsymbol{\sigma}$ be a more-than-proper equilibrium of $\Gamma$. Clearly then, for every $\varepsilon > 0$, and every $\delta \in (0, 1)$, there exists a $(\epsilon\delta)$-more-than-proper equilibrium such that $\boldsymbol{\sigma} = \lim_{\varepsilon \to 0} \boldsymbol{\sigma}^{\varepsilon\delta}$. This follows since an $(\varepsilon\delta)$-more-than-proper equilibrium is $(\varepsilon'\delta)$-more-than-proper for all $\epsilon' \geq \epsilon$. Choose a $\delta$ smaller than $\varepsilon$ and the result immediately follows.

To show that the set is closed fix a $\delta$. Let $C_{\varepsilon\delta}$ denote the closure of the set of $(\overline{\varepsilon}\delta)$-more-than-proper equilibria, for $\overline{\varepsilon} \in (0, \varepsilon)$. Letting $C_\delta$ denote the set of

$\delta$-more-than-proper equilibria for this particular $\delta$, we have that $C_\delta = \cap_{\varepsilon>0} C_{\varepsilon\delta}$. It follows that

$$MTP(\Gamma) = \bigcap_{\delta>0} C_\delta = \bigcap_{\delta>0} \bigcap_{\varepsilon>0} C_{\varepsilon\delta}$$

The possible emptiness of $MTP(\Gamma)$ stems from intersecting over all $\delta \in (0,1)$. It can be shown by a slight adaption of Theorem 17 part 5, that the set $C_\delta$ is non-empty for all $\delta \in (0,1)$. $\square$

## 4.8 Truly Proper Equilibria

Despite being a refinement of properness (see Theorem 17 part 3 below), the attentive reader will have noticed that more-than-properness does not in fact cover the set of all reasonably constrained games.

To be precise, Myerson (1978) constrains, via an upper bound, the weights that *more costly* mistakes receive, but he does not insist that these constraints bind. More-than-properness expands the set of constraints, and insists that the equilibrium is robust for all these constraints. However, the set of reasonable constraints required by more-than-properness is not exhaustive, as it only it insists on constant relative trembling, and only requires robustness to everybody simultaneously trembling according to the same constant weighting scheme. An equilibrium will be referred to as *truly proper* if it holds for *all* weighting schemes that do not violate dominance in the complement of the best response set, and not just the constant weighting schemes used in more-than-properness.

Before presenting the formal statement we introduce some additional notation. Let $m_i$ denote the number of pure strategies of player $i$, i.e. $m_i := |S_i|$. Define $m := \sum_{i=1}^n m_i$. Let $\hat{m}_i$ denote the number of equivalent strategies (as in defintion 27) of player $i$. Also define $\hat{m} := \sum_{i=1}^n \hat{m}_i$. For each player $i$, let $\underline{\delta}^i$ denote an $(\hat{m}_i - 2)$-dimensional vector where each component is between 0 and 1. That is $\underline{\delta}^i := (\delta_i^1, \ldots, \delta_i^{\hat{m}_i-2}) \in \mathbb{R}^{\hat{m}_i-2}$, and represents a profile of relative trembling discount factors for player $i$. Finally, let $\overline{\delta}$ denote a vector of relative trembling discount factors for the game $\Gamma$, that is $\overline{\delta} := (\overline{\delta}^1, \ldots, \overline{\delta}^n) \in \mathbb{R}^{\sum_{i=1}^{n}(\hat{m}_i - 2)} = \mathbb{R}^{(\hat{m}-2n)}$.

**Definition 31.** Let $\Gamma = (S_1, \ldots, S_N, u_1, \ldots, u_n)$ be an $n$-person normal form game. Let $\varepsilon > 0$, and $\overline{\delta} \in \mathbb{R}^{(\hat{m}-2n)}$, and $\boldsymbol{\sigma}^{\varepsilon\overline{\delta}} \in \boldsymbol{\Sigma}$. We call $\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}$ an $(\varepsilon\overline{\delta})$-*truly proper equilibrium* of $\Gamma$ if it is completely mixed and satisfies:

1. if $\quad s_i \in \mathrm{BR}_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}})$,
   then $\quad \sigma_i^{\varepsilon\overline{\delta}}(s_i') \leq \varepsilon \sigma_i^{\varepsilon\overline{\delta}}(s_i)$, for all $i$, for all $s_i' \notin \mathrm{BR}_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}})$.

2. if $\quad c_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}, s_i'') < c_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}, s_i')$, and $\nexists\, \tilde{s}_i \in S_i$ such that
   $c_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}, s_i'') < c_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}, \tilde{s}_i) < c_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}, s_i')$,
   then $\quad \sigma_i^{\varepsilon\overline{\delta}}(s_i') = \frac{1}{|[s_i']|}\delta_i' \sigma_i^{\varepsilon\overline{\delta}}(s_i'')$, all $i$, all $s_i', s_i'' \notin \mathrm{BR}_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}})$, some $\delta_i' \in (0, 1)$

3. if $\quad s_i', s_i'' \notin \mathrm{BR}_i(\boldsymbol{\sigma}^{\varepsilon\overline{\delta}})$, and $s_i', s_i''$ are payoff equivalent
   then $\quad \sigma_i^{\varepsilon\overline{\delta}}(s_i') = \sigma_i^{\varepsilon\overline{\delta}}(s_i'')$, for all $i$.

$\boldsymbol{\sigma}^{\overline{\delta}} \in \boldsymbol{\Sigma}$ is a $\overline{\delta}$-truly proper equilium of $\Gamma$, if $\boldsymbol{\sigma}^{\overline{\delta}}$ is a limit point of a sequence $\left\{\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}\right\}_{\varepsilon\downarrow 0}$ where $\boldsymbol{\sigma}^{\varepsilon\overline{\delta}}$ is an $\varepsilon\delta$-truly proper equilibrium of $\Gamma$ for all $\varepsilon$, for some $\overline{\delta} \in (0, 1)^{(\hat{m}-2n)}$.

Finally we say that $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$ is a *truly proper equilibrium* of $\Gamma$, if $\boldsymbol{\sigma}$ is a $\overline{\delta}$-proper equilibrium of $\Gamma$ for all $\overline{\delta} \in (0, 1)^{(\hat{m}-2n)}$.

*Remark* 9. It is immediately apparent that truly proper equilibria permit a finer selection criterion to more-than-proper equilibria. It should be clear that more-than-properness only has “bite” for games where at least one player has 3 or more pure strategies. Similarly, true properness can only strictly refine the set of more-than-proper equilibria for games where at least one player has 4 or more strategies. For games where no player has 4 or more pure strategies, the two coincide.

*Remark* 10. Note that true properness is a much stronger refinement that more-than-properness in the sense that it requires player $i$ to rank his pure strategies in the same manner for all relative trembling profiles $\underline{\delta}^j$ that all other players $j$ may have. More than properness only requires robustness to each player relatively trembling in the same manner at the same time. True properness requires much more.

*Remark* 11. If more-than-properness is analogous to requiring robustness to the set of all ($\beta\delta$)-discounting model of Phelps and Pollak (1968), then truly proper equilibria are analogous to the set of all hyperbolic discounting models where $U(c_0, c_1, \dots, c_T) = \sum_{t=0}^{T}(\prod_{0\leq s\leq t}\beta_s)u(c_t)$, where $\beta_0 = 1$, and $\beta_t \in [0,1]$ for all $t \in \{1, \dots, T\}$. In other words there is no need for tomorrow to be as relatively less important to today, as today was relatively less important to yesterday. Rather the only condition that must be satisfied is that a day in the future is never more important than any preceding day.

**Lemma 13.** *When they exist, the set of truly proper equilibria, $TP(\Gamma)$, is a closed subset of $MTP(\Gamma)$. Though possibly empty.*

*Proof.* It is not hard to see that more-than-proper equilibria are merely a special case of truly proper equilibria where robustness is only required for the set $\left\{\overline{\delta}\,|\,\overline{\delta} = (\delta, \dots, \delta) \in \mathbb{R}^{\hat{m}-2}, \forall\, \delta \in (0,1)\right\}$. Possible emptiness of $TP(\Gamma)$ is then immediate. Showing the closure of $TP(\Gamma)$ is notationally far more cumbersome than that of lemma 12 but follows along similar lines. It is therefore omitted. □

## 4.9 The Food Chain Theorem

This section categorizes our three new strategy perturbation-based refinements in the respective "food chain" of existing refinements for normal form games.

**Theorem 17.** *1. Every reasonable equilibrium is perfect.*

*2. Every proper equilibrium is reasonable.*

*3. Every more-than-proper equilibrium is proper.*

*4. Every truly proper equilibrium is more-than-proper.*

5. *Every strictly proper equilibrium is truly proper.*

*Proof.* 1. This was shown in Lemma 11.

2. This was shown in Lemma 11.

3. This was shown in Lemma 12

4. This was shown in Lemma 13

5. We will show that every strictly proper equilibrium is more-than-proper. This is for notational ease only. Showing that strict properness implies true properness is proved along very similar lines, though is notationally far messier.

   Assume $\boldsymbol{\sigma}$ is a strictly proper equilibrium of $\Gamma$. Let $O_{\hat{\eta}}$ be as in definition 25. For $\eta \in O_{\hat{\eta}}$ let $\boldsymbol{\sigma}(\eta)$ be as in definition 26.

   Let $\delta \in (0,1)$, and let $m$ be defined as before. We define $\epsilon$ and $U$ as follows:

$$0 < \epsilon < \min_{\substack{s_i \in S_i \\ i \in N}} \hat{\eta}_i(s_i)$$

$$U := \left\{ \eta \in O_{\hat{\eta}} \,\middle|\, \epsilon\delta^{m-1} \leq \eta_i(s_i) \leq \epsilon, \text{ for all } i, \text{ for all } s_i \right\}$$

   Now define the correspondence $G$ from $\boldsymbol{\Sigma}$ to $U$ as follows:

$$\begin{aligned} G(\boldsymbol{\sigma}) := \Big\{ \; \eta \in U \;\Big|\; & (1) \text{ for all } i, \text{ for all } s_i' \notin \mathrm{BR}_i(\boldsymbol{\sigma}) \\ & \quad \text{if} \quad s_i \in \mathrm{BR}_i(\boldsymbol{\sigma}), \\ & \quad \text{then} \quad \eta_i(s_i') \leq \varepsilon \eta_i(s_i), \\ & \quad \text{and} \\ & (2) \text{ for all } i, \text{ for all } s_i' \sim s_i'' \\ & \quad \text{then} \quad \eta_i(s_i') = \eta_i(s_i'') \\ & (3) \text{ for all } i, \text{ for all } s_i', s_i'' \notin \mathrm{BR}_i(\boldsymbol{\sigma}) \\ & \quad \text{if} \quad c_i(\boldsymbol{\sigma}, s_i'') < c_i(\boldsymbol{\sigma}, s_i'), \text{ and } \nexists\, \tilde{s}_i \in S_i \\ & \quad \text{such that} \\ & \quad c_i(\boldsymbol{\sigma}, s_i'') < c_i(\boldsymbol{\sigma}, \tilde{s}_i) < c_i(\boldsymbol{\sigma}, s_i'), \\ & \quad \text{then} \quad \eta_i(s_i') = \tfrac{1}{|[s_i']|} \delta \eta_i(s_i'') \Big\} \end{aligned}$$

It is clear that $G(\boldsymbol{\sigma})$ is a compact and convex set for each $\boldsymbol{\sigma} \in \boldsymbol{\Sigma}$. The upper-hemicontinuity of $G$ follows from the continuity of each player's utility function. To show that $G$ is nonempty we proceed as follows:

For a given profile $\boldsymbol{\sigma}$, each $s_i$ belongs to an equivalence class $[s_i]$. We define a slightly modified version of the count function given in equation 4.3 as follows: For each strategy profile $\boldsymbol{\sigma}$, let

$$\hat{c}(\boldsymbol{\sigma}, [\hat{s}_i]) := \#\Big\{[s_i] \,|\, u_i(\boldsymbol{\sigma}\backslash[s_i]) > u_i(\boldsymbol{\sigma}\backslash[\hat{s}_i])\Big\}$$

where $(\boldsymbol{\sigma}\backslash[s_i])$ denotes the profile $\boldsymbol{\sigma}$, with $\sigma_i$ replaced with any pure strategy of player $i$ that is payoff equivalent (against $\boldsymbol{\sigma}$) to $s_i$. Thus the modified count function, $\hat{c}(\boldsymbol{\sigma}, \cdot)$ ranks the equivalence classes for a player relative to a given strategy profile, $\boldsymbol{\sigma}$. Define

$$\eta_i(s_i) = \Big(\frac{1}{\prod_{\{[s_i'']:\hat{c}(\boldsymbol{\sigma},[s_i''])\leq\hat{c}(\boldsymbol{\sigma},[s_i])\}} |[s_i'']|}\Big)\frac{\varepsilon\delta^{\hat{c}(\boldsymbol{\sigma},[s_i])}}{\varepsilon\sum_{[s_i']}\delta^{\hat{c}(\boldsymbol{\sigma},[s_i'])}}$$

Clearly $\eta_i \in G_i(\boldsymbol{\sigma})$, where $G_i$ is the component of $G$ associated with player $i$, i.e. $G = G_1 \times \cdots \times G_n$.

Now define a correspondence $F$ from $U$ to $U$ by $F(\eta) = G(\boldsymbol{\sigma}(\eta))$, where $\boldsymbol{\sigma}(\eta) \in E(\Gamma^\eta)$. It is clear that $F$ satisfies the sufficient conditions of the Kakutani (1941) fixed point theorem, and therefore has a fixed point in $U$. Let $\eta^{\epsilon\delta}$ denote this fixed point. To minimize notation, we will denote $\boldsymbol{\sigma}(\eta^{\epsilon\delta})$ by $\boldsymbol{\sigma}^{\varepsilon\delta}$.

We thus have constructed the following:

$$\begin{aligned} &\text{if } s_i \in \mathrm{BR}_i(\boldsymbol{\sigma}), \\ &\text{then } \eta_i^{\varepsilon\delta}(s_i') \leq \varepsilon\eta_i^{\varepsilon\delta}(s_i), \text{for all } i, \text{ for all } s_i' \notin \mathrm{BR}_i(\boldsymbol{\sigma}). \end{aligned} \tag{4.4}$$

$$
\begin{aligned}
&\text{if } c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s_i'') < c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s_i'),\\
&\text{and } \nexists\, \tilde{s}_i \in S_i \text{ such that } c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s_i'') < c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, \tilde{s}_i) < c_i(\boldsymbol{\sigma}^{\varepsilon\delta}, s_i'),\\
&\text{then } \eta_i^{\varepsilon\delta}(s_i') = \frac{1}{|[s_i']|}\delta\eta_i^{\varepsilon\delta}(s_i''), \text{ for all } i, \text{ for all } s_i', s_i'' \notin \mathrm{BR}_i(\boldsymbol{\sigma}^{\varepsilon\delta}).
\end{aligned} \tag{4.5}
$$

Now, since $\boldsymbol{\sigma}^{\varepsilon\delta}$ is an equilibrium of the constrained game $(\Gamma^{\eta^{\epsilon\delta}})$, the players must abide by the constraints. Hence we must have

$$
\sigma_i^{\varepsilon\delta}(s_i) \geq \eta_i^{\epsilon\delta}(s_i), \text{ for all } i, \text{ and all } s_i \in S_i \tag{4.6}
$$

Furthermore, it can trivially be proved (see for example Van Damme (1996) Lemma 2.2.2 page 27) that the constraints on any pure strategy that is not a best response must bind in a constrained game, so

$$
\begin{aligned}
&\text{if } \quad u_i(\boldsymbol{\sigma}^{\varepsilon\delta}\backslash s_i') < u_i(\boldsymbol{\sigma}^{\varepsilon\delta}\backslash s_i'),\\
&\text{then } \quad \sigma_i^{\varepsilon\delta}(s_i') = \eta_i^{\epsilon\delta}(s_i'), \text{ for all } i, \text{ for all } s_i', s_i''.
\end{aligned} \tag{4.7}
$$

It is clear that by combining the equations from (4.4) - (4.7), we have that $\boldsymbol{\sigma}^{\varepsilon\delta}$ is an $\epsilon\delta$-more-than-proper equilibrium of $\Gamma$. For $\epsilon$ very small, we have that $\boldsymbol{\sigma}^{\varepsilon\delta}$ is close to $\boldsymbol{\sigma}$. As such $\boldsymbol{\sigma}$ is a $\delta$-proper equilibrium of $\Gamma$.

Finally we note that our choice of $\delta$ was arbitrary, and thus $\boldsymbol{\sigma}$ is a $\delta$-more-than-proper equilibrium uniformly in $\delta$. $\boldsymbol{\sigma}$ is therefore more-than-proper.

□

## 4.10 Conclusion

This paper extends the literature on rational behaviour in multiplayer non-cooperative games. Our main goal is to reexamine the notion of what are considered reasonable beliefs that players may hold in equilibrium. Specifying reasonable off-path equilibrium beliefs has always been done by looking at the limit of fully mixed strategies. The question we sought to answer is a standard one: What kind of trembles are sensible?

We approach this problem from a different perspective by flipping the question on its head, instead asking: What is the set of trembles that don't make intuitive sense? We assume perturbations of pure strategies that have full support, and insist on the quite natural requirement that weights on pure strategies are ordered according to conditional dominance. Any strategy profile in which any player assigns more weight to an inferior strategy is deemed unreasonable and is rejected.

The set of perturbations not violating conditional dominance are referred to as *reasonable*. An equilibrium is reasonable if it is robust to one sequence of perturbations from the reasonable set. An equilibrium is referred to as more than proper if it is robust to a large yet easily characterized set of perturbations from the reasonable set. An equilibrium is referred to as truly proper if it is robust to all the perturbations from the reasonable set.

It is widely accepted that equilibria should be insensitive to some strategy perturbations - as such requiring robustness to some (if not all such perturbations is essential). It is not so widely accepted that strategy perturbations should be thought of as limiting beliefs rather than a mere technical trick. We argue that if any behavioral interpretation is assigned to trembles, then reasonableness is surely a necessary condition for an equilibrium to be self enforcing. Modeling is a key component of economics, and while models break down in all sorts of ways, possibly (surely?) in a few cases, some modes of breakdown are more plausible than others. So requiring robustness to these modes seems all the more important. Furthermore, requiring robustness to certain modes of breakdown can at times be seen as odd.

Potential extensions and applications of our refinements abound. Neary (2008) examines what happens when there is common knowledge of reasonable trembling but players have uncertainty over what type of *reasonable trembler* their opponents are. Such an assumption selects the *reasonably rationalizable* strategies of a game. Another interesting line of research could be to weaken the restrictions in learning-theoretic foundations of equilibrium (see Fudenberg and Levine (1993)), to a point where players need not even agree ex-post with on path beliefs, they just need have *similar* (and of course reasonable) beliefs. Recently, refinements of equilibria have been applied in many varied subfields of microeconomics: to the networks literature, see Calva-Armengol and Rahmi (2007) for example; to voting games, see Sinopoli (2000) and Sinopoli, Dutta, and Laslier (2006); and of course there is a large literature on refinements as applied to signaling games, see Banks and Sobel (1987) and Cho and Kreps (1987). Our refinements are tractable and supply another method for refinements in such games.

Finally, while this paper has focused mostly on reasonable equilibria, various referees have expressed interest in both more-than-properness and true-properness but are concerned by the lack of existence of both. An interesting avenue to explore would be to extend both point-valued concepts to set-valued versions, and to check whether these sets satisfy the desired requirements of stability of Kohlberg and Mertens (1986), and Mertens (1989, 1991). Hillas (1990) also looks at set-valued solutions by requiring robustness to perturbations of the best-response correspondence. One could modify his analysis by restricting the set of allowable perturbations to the best-response correspondence and constrain them to be reasonable. This is left to future research.

# Bibliography


Arthur, W. B. (1989): "Competing Technologies, Increasing Returns, and Lock-In by Historical Events," *Economic Journal*, 99(394), 116–31.

Austen-Smith, D., and R. G. Fryer (2005): "An Economic Analysis of "Acting White"," *The Quarterly Journal of Economics*, 120(2), pp. 551–583.

Banks, J., and J. Sobel (1987): "Equilibrium Selection in Signaling Games," *Econometrica*, 55(3), 647–661.

Bergin, J., and B. L. Lipman (1996): "Evolution with State-Dependent Mutations," *Econometrica*, 64(4), 943–956.

Blume, L. E. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5(3), 387 – 424.

Bramoulle, Y. (2007): "Anti-coordination and social interactions," *Games and Economic Behavior*, 58(1), 30–49.

Bramoulle, Y., and R. Kranton (2007): "Public goods in networks," *Journal of Economic Theory*, 135(1), 478–494.

Calva-Armengol, A., and Rahmi (2007): "Pairwise-stability and Nash equilibria in network formation," *International Journal of Game Theory*.

Canals, J., and F. Vega-Redondo (1998): "Multi-level evolution in population games," *International Journal of Game Theory*, 27(1), 21–35.

Cho, I.-K., and D. M. Kreps (1987): "Signaling Games and Stable Equilibria," *The Quarterly Journal of Economics*, 102(2), 179–221.

Cressman, R., J. Garay, and J. Hofbauer (2001): "Evolutionary Stability Concepts for N-species Frequency-dependent Interactions," *Journal of Theoretical Biology*, 211(1), 1 – 10.

Dasgupta, P., and E. Maskin (1986): "The Existence of Equilibrium in Discontinuous Economic Games, I: Theory," *The Review of Economic Studies*, 53(1), pp. 1–26.

Dekel, E., J. C. Ely, and O. Yilankaya (2007): "Evolution of Preferences," *Review of Economic Studies*, 74(3), 685–704.

Eaves, B. C. (1973): "Polymatrix Games with Joint Constraints," *SIAM Journal on Applied Mathematics*, 24(3), pp. 418–423.

Ellison, G. (1993): "Learning, Local Interaction, and Coordination," *Econometrica*, 61(5), 1047–1071.

——— (2000): "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution," *Review of Economic Studies*, 67(1), 17–45.

Ely, J. (2002): "Local Conventions," *Advances in Theoretical Economics*, 2(1).

Farrell, J., and G. Saloner (1985): "Standardization, Compatibility, and Innovation," *RAND Journal of Economics*, 16(1), 70–83.

Foster, D., and P. Young (1990): "Stochastic evolutionary game dynamics," *Theoretical Population biology*, 38, 219–232.

Freidlin, M. I., and A. D. Wentzell (1998): *Random Perturbations of Dynamical Systems (Grundlehren der mathematischen Wissenschaften)*. Springer Verlag.

Friedman, D. (1998): "On economic applications of evolutionary game theory," *Journal of Evolutionary Economics*, 8(1), 15–43.

Fudenberg, D., and D. K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61(3), 523–545.

——— (1998): *The Theory of Learning in Games (Economic Learning and Social Evolution)*. The MIT Press.

Fudenberg, D., and E. Maskin (1986): "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information," *Econometrica*, 54(3), 533–554.

Fudenberg, D., and J. Tirole (1991): *Game Theory*. MIT Press.

Galeotti, A., S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv (2009): "Network Games," *Review of Economic Studies*.

Goyal, S. (2007): *Connections: Connections: An Introduction to the Economics of Networks*. Princeton University Press.

Harsanyi, J. C., and R. Selten (1988): *A General Theory of Equilibrium Selection in Games*. MIT Press.

HILLAS, J. (1990): "On the Definition of the Strategic Stability of Equilibria," *Econometrica*, 58(6), 1365–1390.

HOFBAUER, J., AND K. SIGMUND (1998): *Evolutionary Games and Population Dynamics*. Cambridge University Press.

HOWSON, JOSEPH T., J. (1972): "Equilibria of Polymatrix Games," *Management Science*, 18(5), pp. 312–318.

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

JACKSON, M. O., AND A. WATTS (2002): "The Evolution of Social and Economic Networks," *Journal of Economic Theory*, 106(2), 265–295.

——— (2008): "Social games: Matching and the play of finitely repeated games," *Games and Economic Behavior*, In Press, Corrected Proof, –.

JANOVSKAYA, E. B. (1968): "Equilibrium Points in Polymatrix Games," *(in Russian), Latvian Mathematical Journal*.

JURADO, I. G., AND J. M. P. SÁNCHEZ (1990): "A remark on Myerson's concept of proper equilibrium," *International Journal of Game Theory*, 19(1), 11–16.

KAKUTANI, S. (1941): "A Generalization of Brouwer's Fixed Point Theorem," *Duke Mathematical Journal*, 8(3), 457–459.

KANDORI, M., G. J. MAILATH, AND R. ROB (1993): "Learning, Mutation, and Long Run Equilibria in Games," *Econometrica*, 61(1), 29–56.

KARLIN, S., AND H. M. TAYLOR (1975): *A First Course in Stochastic Processes, Second Edition*. Academic Press, 2 edn.

KATZ, M. L., AND C. SHAPIRO (1985): "Network Externalities, Competition, and Compatibility," *American Economic Review*, 75(3), 424–40.

KEARNS, M., S. SURI, AND N. MONTFORT (2006): "An Experimental Study of the Coloring Problem on Human Subject Networks," *Science*, 313(5788), 824–827.

KIM, Y.-G., AND J. SOBEL (1995): "An Evolutionary Approach to Pre-Play Communication," *Econometrica*, 63(5), 1181–1193.

KOHLBERG, E., AND J.-F. MERTENS (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54(5), 1003–1037.

KREPS, D. M., AND R. WILSON (1982): "Sequential Equilibria," *Econometrica*, 50(4), 863–894.

Lewis, D. K. (1969): *Convention: a Philosophical Study*. Harvard University Press.

Manea, M. (2009): "Bargaining in Stationary Networks," Discussion paper, Harvard.

Matsui, A. (1991): "Cheap-talk and cooperation in a society," *Journal of Economic Theory*, 54(2), 245–258.

McCubbins, M. D., R. Paturi, and N. Weller (2008): "Networked Coordination: The Effect of Network Structure on Human Subjects' Ability to Coordinate," *SSRN eLibrary*.

Mertens, J.-F. (1989): "Stable Equilibria: A Reformulation Part I. Definition and Basic Properties," *Mathematics of Operations Research*, 14(4), 575–625.

——— (1991): "Stable Equilibria: A Reformulation. Part II. Discussion of the Definition, and Further Results," *Mathematics of Operations Research*, 16(4), 694–753.

Morris, S. (2000): "Contagion," *Review of Economic Studies*, 67(1), 57–78.

Myerson, R. B. (1978): "Refinements of the Nash equilibrium concept," *International Journal of Game Theory*, 7(2), 73–80.

Neary, P. R. (2008): "Reasonable Beliefs in Games," Discussion paper, University of California San Diego, Department of Economics.

——— (2010a): "Competing Conventions," Discussion paper, University of California San Diego.

——— (2010b): "Competing Conventions in Networked Societies," Discussion paper, University of California San Diego.

——— (2010c): "Multiple Group Games," Discussion paper, University of California San Diego, Department of Economics.

Oechssler, J. (1997): "Decentralization and the coordination problem," *Journal of Economic Behavior and Organization*, 32(1), 119 – 135.

——— (1999): "Competition Among Conventions," *Computational & Mathematical Organization Theory*, 5(1), 31–44.

Ok, E. A., and F. Vega-Redondo (2001): "On the Evolution of Individualistic Preferences: An Incomplete Information Scenario," *Journal of Economic Theory*, 97(2), 231–254.

Okada, A. (1981): "On stability of perfect equilibrium points," *International Journal of Game Theory*, 10.

Peski, M. (2010): "Generalized risk-dominance and asymmetric dynamics," *Journal of Economic Theory*, 145(1), 216–248.

Phelps, E. S., and R. A. Pollak (1968): "On Second-Best National Saving and Game-Equilibrium Growth," *The Review of Economic Studies*, 35(2), 185–199.

Polanski, A. (2007): "Bilateral bargaining in networks," *Journal of Economic Theory*, 134(1), 557–565.

Rapoport, A., and M. J. Guyer (1966): "A Taxonomy of 2 x 2 Games," *General Systems*, 11, 203–214.

Robinson, D. R., and D. J. Goforth (2005): *The Topology of the 2x2 Games: A New Periodic Table*. Routledge.

Robson, A. J., and F. Vega-Redondo (1996): "Efficient Equilibrium Selection in Evolutionary Games with Random Matching," *Journal of Economic Theory*, 70(1), 65 – 92.

Rubinstein, A. (1979): "Equilibrium in supergames with the overtaking criterion," *Journal of Economic Theory*, 21(1), 1–9.

Samuelson, L. (1994): "Stochastic Stability in Games with Alternative Best Replies," *Journal of Economic Theory*, 64(1), 35–65.

Samuelson, L., and J. Zhang (1992): "Evolutionary stability in asymmetric games," *Journal of Economic Theory*, 57(2), 363–391.

Sandholm, W. H. (2011): *Population Games and Evolutionary Dynamics (Economic Learning and Social Evolution)*. The MIT Press.

Schelling, T. C. (1969): "Models of Segregation," *The American Economic Review*, 59(2), pp. 488–493.

——— (1971): "Dynamic models of segregation," *Journal of Mathematical Sociology*, 1, 143–186.

Selten, R. (1975): "Reexamination of the perfectness concept for equilibrium points in extensive games," *International Journal of Game Theory*, 4(1), 25–55.

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

Sinopoli, F., B. Dutta, and J.-F. Laslier (2006): "Approval voting: three examples," *International Journal of Game Theory*, 35(1), 27–38.

SINOPOLI, F. D. (2000): "Sophisticated voting and equilibrium refinements under plurality rule," *Social Choice and Welfare*, 17(4), 655–672.

SMITH, J. M. (1982): *Evolution and the Theory of Games*. Cambridge University Press, 1st edition edn.

SMITH, J. M., AND G. R. PRICE (1973): "The Logic of Animal Conflict," *Nature*, 246(5427), 15–18.

VAN DAMME, E. (1984): "A Relation between Perfect Equilibria in Extensive Form Games and Proper Equilibria in Normal Form Games," *International Journal of Game Theory*, 13.

——— (1996): *Stability and Perfection of Nash Equilbria*. Springer-Verlag Berlin Heidelberg.

WEIBULL, J. W. (1995): *Evolutionary Game Theory*. Cambridge: MIT Press.

YOUNG, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.

——— (1996): "The Economics of Convention," *The Journal of Economic Perspectives*, 10(2), 105–122.

——— (2001): *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton University Press.

——— (2005): *Strategic Learning and Its Limits (Arne Ryde Memorial Lectures Sereis)*. Oxford University Press, USA.