Essays in Evolutionary Game Theory

Ge Jiang

A thesis submitted for the degree of Doctor of Philosophy

Department of Economics

University of Essex

July 2016

Statement of Co-authorship

PhD candidate: Ge Jiang

Paper Title: Local Interactions under Switching Costs (Chapter 1 in the thesis)

List of authors: Ge Jiang, Simon Weidenholzer

Ge Jiang's contribution in the paper includes: (1) The model development (Section 2); (2) The two-neighbor interaction. (Section 3.1); (3) The 2k-neighbor interaction and global interactions (Section 3.2, including Lemma 1 and Proposition 2); (4) Drafting the article. Simon Weidenholzer's contribution in the paper includes: (1) Corollary 3; (2) Corrections and recommendations to improve the article.

Acknowledgments

The completion of this thesis has led to the accumulation of enormous debts, to a great many people, that I will probably never be able to repay. First and foremost, I would like to thank my supervisor Simon Weidenholzer, and advisor, Andrea Galeotti, for their ongoing support, ´ their patience and their insights, without which I would have been lost. I have also benefited immoderately from discussions with, and comments from, other faculty and colleagues in Department of Economics in University of Essex. I am also very indebted to Zhiwei Cui for his advice and substantive comments on my work. Audiences at seminars and anonymous referees have provided useful feedback on several parts of the thesis.

Last, but by no means least, my family and friends deserve special thanks for putting up with me with such grace and good humour throughout this long journey.

Abstract

This thesis contains three essays in evolutionary game theory.

In the first chapter, we study the impact of switching costs on the long run outcome in 2×2 coordination games played in the circular city model of local interactions. We find that for low levels of switching costs, the risk dominant convention is the unique long run equilibrium. For intermediate levels of switching costs the set of long run equilibria contains the risk dominant convention but may also contain conventions that are not risk dominant. For high levels of switching costs also nonmonomorphic states will be included in the set of LRE.

We study the impact of location heterogeneity on neighborhood segregation in the one-dimensional Schelling residential model in the second chapter. We model location heterogeneity by introducing an advantageous node, in which a player's utility is impartial to the composition of her neighborhood. We find that when every player interacts with two neighbors, one advantageous node in the circular city will lead to a result that segregation is no longer the unique LRE. When players interact with more neighbors, more advantageous nodes are necessary to obtain the same result.

In the third chapter, we consider a model of social coordination and network formation, where players of two groups play a 2×2 coordination game when connected. Players in one group actively decide on whom they play with and on the action in the game, while players in the other group decide on the action in the game only. We find that if either group's population size is small in comparison to the linking restriction, all players will choose the risk dominant equilibrium, while when both groups are sufficiently large in population, the players of two groups will coordinate on the payoff dominant action.

Contents

Local Interactions under Switching Costs 6

1 Introduction **2**

2 The model **7**

3 The role of switching costs **10**
3.1 Two-neighbor interaction 12
3.2 $2k$-neighbor interaction and global interactions 14

4 Conclusion **22**

Schelling's Model Revisited: From Segregation to Integration 26

1 Introduction **27**

2 Literature review **29**

3 The model **32**
3.1 Basic Setup 32
3.2 The dynamic process 33
3.3 Review of techniques 34

4 Location heterogeneity in the two-neighbor residential model **35**

5 Location heterogeneity in the $2k$-neighbor residential model **42**
5.1 The $2k$-neighbor model without advantageous node 42
5.2 Multiple advantageous nodes in the $2k$-neighbor model 46

6 Simulations **52**

7 Conclusion **54**

Social Coordination and Network Formation in Two Groups 57

1 Introduction **58**

2 Literature Review **60**

3 The Model **62**
3.1 Review of Techniques 66

4 Coordination in Two Groups **68**
4.1 Long-run Equilibria 76

5 Conclusion **78**

Local Interactions under Switching Costs

Abstract

We study the impact of switching costs on the long run outcome in 2×2 coordination games played in the circular city model of local interactions. For low levels of switching costs the predictions are in line with the previous literature and the risk dominant convention is the unique long run equilibrium. For intermediate levels of switching costs the set of long run equilibria still contains the risk dominant convention but may also contain conventions that are not risk dominant. The set of long run equilibria may further be non-monotonic in the level of switching costs, i.e. as switching costs increase the prediction that the risk dominant convention is unique long run equilibrium and the prediction that both conventions are long run equilibria alternate. Finally, for high levels of switching costs also non-monomorphic states will be included in the set of long run equilibria.

1 Introduction

It is often costly to switch to a different technology or adopt a new social norm. For instance, switching from Windows to Apple requires not only getting familiarized to the new system but also moving files from one computer to the other. Further examples of switching costs include communicating one's new telephone number when switching providers in telecommunication or buying new tools when switching from inch screws to metric screws.

The present paper aims to understand the role of switching costs on long run technology choice and the emergence of conventions. Since agents are better off when interacting with somebody who uses the same operating system, telecommunication provider, or industry standard, these situations typically give rise to coordination games. A wide range of models, starting with the seminal works of Kandori, Mailath and Rob (1993) and Young (1993), have analyzed settings where a population of boundedly rational players decide on their actions in such coordination games using simple heuristics.[1] The message that emerges from these discussions is that, when players use best response learning, risk dominant strategies -that perform well against mixed strategy profiles will emerge in the long run, even in the presence of payoff dominant strategies. In the context of the above examples, this implies that populations do not necessarily end up with technologies which maximize social welfare.

Norman (2009) has already analyzed the role of switching costs in a global interactions setting where everybody interacts with everybody else. In the global setting switching costs turned out to influence the speed at which the population approaches the long run equilibrium (LRE). The long run prediction remain unaffected, though. Quite frequently interactions are, however, local in nature, with interaction partners corresponding to family members, friends, or work colleagues. For instance, in the above examples on switching operating systems or telecommunication providers it is

[1] See Weidenholzer (2010) for a survey of the literature.

typically the case that this decision will to a larger degree be influenced by one's contacts or collaborators than by the overall distribution of technologies in the society. We capture such local interactions by considering a model akin to the one proposed by Ellison (1993) where the agents are arranged around a circle and interact with their neighbors only. We focus on a setting where one strategy is risk dominant and the other strategy may or may not be payoff dominant. This allows us to analyze circumstances under which strategies that are neither payoff- nor risk- dominant are selected. When determining which strategy to use the players play a best response to the distribution of play in their neighborhood in the previous period taking into account that switching strategies incurs a cost. In addition, choices are perturbed by occasional uniform (across agents and time) mistakes.

We find that low levels of switching costs do not change the predictions of the model as compared to the standard model without switching costs. The risk dominant strategy is still able to spread contagiously, starting from a small cluster and eventually taking over the whole population. However, for larger switching costs risk dominant strategies may no longer spread contagiously and non-monomorphic states, where different strategies coexist, become absorbing. The reason is that a player at the boundary of a risk dominant cluster will not switch under sufficiently high switching costs. It is possible to move among all of these non-monomorphic absorbing states via a chain of single mutations. Transitions from different states to each others are, thus, characterized by step-by-step evolution as outlined in Ellison (2000).

The question which state will be LRE essentially boils down to how difficult the set of non-monomorphic states is to access from the two monomorphic states. Interestingly, if agents only interact with a few neighbors, there may exist a range of parameters where alongside the risk dominant convention also non-risk dominant conventions are LRE. Thus, switching costs may lead to the model's prediction no longer being unique. The reason behind this phenomenon is that under the uniform noise approach the number of mutations required to move from a convention to the set of non-monomorphic absorbing states is measured in integers. Especially if agents only interact with a few neighbors, it may happen that the number of mistakes required to access the set of non-monomorphic absorbing states from the risk dominant convention equals the number of mistakes required to access this set from the non-risk dominant convention.

Perhaps even more interestingly, also owing to the fact the mutations are measured in integers, the prediction might be non-monotonic in the level of switching costs. That is, the prediction that the risk dominant convention is selected and the prediction that both conventions are selected alternate as switching costs increase. This curiosity is caused by i) the stepwise nature of rounding up and ii) by the fact that the number of mistakes required to leave the risk dominant convention and the number of mistakes required to leave the non-risk dominant convention only differ by a constant.

Finally, for very high levels of switching costs no player will switch in the absence of noise even if all neighbors choose the other strategy. Thus, all states are absorbing and can be connected via a chain of single mutations. Consequently, all absorbing states turn out to be LRE.

For large interaction neighborhoods the integer problem ceases to have impact and the risk dominant convention remains as unique LRE. In particular, this holds true if every agent interacts with every other agent and a sufficiently large population, thus, reconciling our results with those of Norman (2009).

If one takes the model's predictions at face value, it may contribute to our understanding of the emergence and survival of (risk dominated) technology standards or norms. If the risk dominant strategy is not payoff dominant, then the presence of switching costs implies that payoff dominant conventions will be observed with positive probability in the long run. Switching cost might, thus, be welfare improving. If, however, a strategy is both risk- and payoff- dominant the presence of switching costs may lead to (risk- and payoff-) dominated strategies surviving in the long run. Switching costs and local interactions may, thus, also explain why inefficient technology standards or norms survive in the long run.

A more pessimistic reading of our results is that the local interaction model may lose traction in the presence of switching costs as it can no longer give a clear cut prediction. This is expressed by the non-uniqueness of the long run prediction but even more aggravated by the non-monotonicity of the prediction. While the risk dominant convention ceases to be unique LRE for high enough switching costs it might be again unique LRE for even higher switching costs. This is bad news since the circular city model of local interactions has some otherwise nice features as compared to the global model: i) It was observed by Ellison (1993) that in contrast to the global interaction model of Kandori, Mailath, and Rob (1993) it features a high speed of convergence. ii) Lee, Szeidl, and Valentinyi (2003) have shown that it is immune against the Bergin and Lipman (1996) critique. iii) Weidenholzer (2012) has shown that it is robust to the addition and deletion of dominated strategies, a test which Kim and Wong (2010) have shown the global model fails.

The paper closest related to our work is Norman (2009) who studies switching costs in the context of a global interactions model. As already observed by Kandori, Mailath, and Rob (1993) a major drawback of the global interactions model lies in its low speed of convergence. Under global interactions the number of mistakes required to move from one convention to another turns out to depend on the population size. Thus, in large populations it is questionable whether the long run limit will be observed within any reasonable time horizon.[2] Norman (2009) shows how switching costs might speed up convergence. As in the present paper, the presence of switching costs implies that non-monomorphic states where agents use different actions become

[2] Ellison (1993) pointed out that in the context of local interactions where some strategies might spread contagiously the speed of convergence is independent of the population size and, thus, the LRE might be a reasonable predictor even in large populations.

absorbing. This enables a transition from one convention to another by first accessing the class of non-monomorphic states and then moving through this class via a chain of single mutations to the other convention. Under switching cost the step from one convention to the set of non-monomorphic states is typically smaller than the direct step from that convention to the other. Consequently, switching costs may speed up the convergence to the long run prediction.

While the present paper adds to the ongoing discussion on learning in coordination games it also contributes to a wider discussion on how far received results in the literature on learning in games are robust to (minor) modifications. For instance, under imitation learning changing the interaction or information structures may result in different predictions in coordination games (see Robson and Vega-Redondo (1996) and Alós-Ferrer and Weidenholzer (2006), (2008)) or prisoner dilemma games (see Eshel, Samuelson, and Shaked (1998), Mengel (2009)). Similarly, in Cournot games a number of contributions have analyzed conditions under which firms converge to the Walrasian state under imitation learning, as predicted by Vega-Redondo (1997). Alós-Ferrer (2004) shows that when agents have memory over the last two periods the Walrasian state is no longer uniquely stochastically stable.[3] Apesteguia, Huck, Oechssler, and Weidenholzer (2010) find that, if there are differences in cost functions, all monomorphic states are absorbing. However, the Walrasian state remains unique LRE if no firm is the uniquely cheapest one, as shown by Tanaka (1999).

The rest of this paper is organized in the following way. Section 2 presents the model and discusses the main techniques used. Section 3 spells out our main results and Section 4 concludes.

[3] Alós-Ferrer and Shi (2012) consider asymmetric memory which turns out to affect equilibrium selection in coordination games but reinforces the stability of the Walrasian state in Cournot games.

2 The model

We consider a population of N agents who are located on a circle, as in Ellison (1993). A given agent i has agents $i-1$ and $i+1$ (mod N) as immediate neighbors. Each agent interacts with her k closest neighbors to the left and to the right of her. We assume $k \leq \frac{N-1}{2}$ to ensure that no agent interacts with herself. Thus, agent i's interactions are confined to the set of players $N(i) = \{i-k, i-k+1 \ldots, i-1, i+1, \ldots, i+k-1, i+k\}$. The agents in the set $N(i)$ are called neighbors of i.

We assume $|N|$ to be odd. This allows us to nest global interactions in our framework by setting $k = \frac{N-1}{2}$.[4]

Each agent i plays a 2×2 coordination game with strategy set $S = \{A, B\}$ against all agents in her neighborhood $N(i)$. We denote by $u(s_i, s_j)$ the payoff agent i with strategy s_i receives when playing against agent j with strategy s_j. We follow Eshel, Samuelson and Shaked (1998) and consider the following (normalized) coordination game.

	A	B
A	α, α	$\beta, 0$
B	$0, \beta$	$1, 1$

We assume $\alpha > 0$ and $\beta < 1$, so that (A, A) and (B, B) are both strict Nash equilibria. Further, we assume $\alpha + \beta > 1$, so that the equilibrium (A, A) is risk dominant in the sense of Harsanyi and Selten (1988), i.e. A is the unique best response to a mixed strategy profile which puts equal probability on A and B. We denote by

$$q^* = \frac{1-\beta}{1+\alpha-\beta}$$

the critical mass put on A in a mixed strategy equilibrium. Risk dominance of the Nash equilibrium (A, A) translates into $q^* < \frac{1}{2}$. Note that if $\alpha > 1$, (A, A) is payoff

[4] The results obtained for local interaction also hold for even populations.

dominant and if $\alpha < 1$, (B, B) is payoff dominant. However, no such assumption on α is made at this stage.

The number of A-players in the population is denoted by $m = \#\{i \in I | s_i = A\}$ and the number of A-players among agent i's neighbors is denoted by $m_i = \#\{j \in N(i) | s_j = A\}$. Accordingly, the number of B-players in the population is given by $N - m$ and the number of B-players in i's interaction set is given by $2k - m_i$.

We denote by $s_i(t)$ the strategy adopted by player i, by $s(t) = (s_1(t), \ldots, s_N(t))$ the profile of strategies adopted by all players, and by

$$s_{-i}(t) = (s_{i-k}(t), \ldots, s_{i-1}(t), s_{i+1}(t), \ldots, s_{i+k}(t))$$

the strategies adopted by all of player i's neighbors in period t. Further, the monomorphic states $(s, s, \ldots, s)$ where all agents adopt the same strategy s are denoted by $\overrightarrow{s}$.

The payoff for player i is given by the average payoff received when interacting with all neighbors.

$$U_i\big(s_i(t), s_{-i}(t)\big) = \frac{1}{2k} \sum_{j \in N(i)} u\big(s_i(t), s_j(t)\big).$$

We consider a myopic best response process with *switching costs*. In each period $t = 1, 2, \ldots$ each agent receives the opportunity to revise her strategy with exogenous probability $\eta \in (0, 1)$.[5] Changing strategies is assumed to be costly. Whenever an agent changes her strategy she is subject to a switching cost. We follow Norman (2009) and consider switching costs c which are independent of the current action choice and enter the payoff function in an additive way. [6] The following function

[5] Thus, we are considering a model of positive inertia where agents may not adjust their strategy every period.

[6] Alternative formulations of switching costs encompass situations where the level of switching costs depends on the current strategy used or on the current level of payoffs.

formalizes this idea

$$c\left(s_i(t), s_i(t+1)\right) = \begin{cases} c & \text{if } s_i(t) \neq s_i(t+1) \\ 0 & \text{if } s_i(t) = s_i(t+1) \end{cases} .$$

When a revision opportunity arises an agent switches to a myopic best response, i.e. she plays a best response to the distribution of play in her neighborhood in the previous period, taking into account the switching costs. More formally, at time $t+1$, when given revision opportunity, player i chooses

$$s_i(t+1) \in \arg\max_{s_i(t+1)\in S} \Big[U\big(s_i(t+1), s_{-i}(t)\big) - c\big(s_i(t), s_i(t+1)\big)\Big].$$

If a player has multiple best replies, it is assumed that she randomly chooses one of them with exogenously given probability. If she does not receive an opportunity to revise her strategy, she chooses $s_i(t+1) = s_i(t)$. Further, with fixed probability $\epsilon > 0$, independent across agents and across time, the agent ignores her prescription and chooses a strategy at random, i.e. she makes a mistake or mutates.

We denote the state space by Ω and a state of the process by ω. The process with mistakes is called *perturbed process*. Under the perturbed process any two states can be reached from each other. Thus, the only absorbing set is the entire state space, implying that the process is ergodic. The unique invariant distribution of this process is denoted by $\mu(\epsilon)$. We are interested in the limit invariant distribution (as the rate of experimentation tends to zero), $\mu^* = \lim_{\varepsilon\to 0}\mu(\varepsilon)$. Such a distribution exists (see Foster and Young (1990), Young (1993), or Ellison (2000)) and is an invariant distribution of the process without mistakes (the so called *unperturbed process*). It gives a stable prediction for the original process, in the sense that for ϵ small enough the play approximates that described by μ^* in the long run. The states in the support of μ^*, are called *Long Run Equilibria (LRE)* or stochastically stable states. The set of LRE is denoted by $\mathcal{S} = \{\omega \in \Omega \mid \mu^*(\omega) > 0\}$. We use a characterization of the set of LRE due to Freidlin and Wentzell (1988).[7] Consider two absorbing sets of states X and Y and let $C(X,Y) > 0$ (referred to as a *transition cost*) denote the minimal number of mutations for a transition from the X to Y. An X-tree is a directed tree such that the set of nodes is the set of all absorbing sets, and the tree is directed into the root X. For a given tree one can calculate the cost as the sum of the costs of transition for each edge. According to Freidlin and Wentzell (1988), a set X is a LRE if and only if it is the root of a minimum cost tree.

3 The role of switching costs

In a first step we will study how switching costs influence the agent's decision to switch strategies. Consider an A-player. She will switch strategies with probability one if her payoff from playing B minus the switching cost strictly exceeds her payoff from remaining an A- player, i.e.

$$\frac{1}{2k}\Big(m_i\alpha + (2k - m_i)\beta\Big) < \frac{1}{2k}\Big(2k - m_i\Big) - c.$$

Rearranging terms yields

$$m_i < 2kq^* - \frac{2kc}{1+\alpha-\beta} := m^A(c,k).$$

An A-player will remain an A-player with certainty whenever $m_i > m^A(c,k)$ and will choose A and B with positive probability if $m_i = m^A(c,k)$. As $m^A(c,k)$ is the minimum number of A-playing neighbors such that keeping A is a unique best response, it cannot be negative.

Likewise, consider a B-player. She will switch strategies with probability one if the payoff from playing A minus the switching cost exceeds her current payoff, which

[7] See Fudenberg and Levine (1998) or Samuelson (1997) for textbook treatments. Ellison (2000) provides an enhanced (and sometimes easier to apply) algorithm for identifying the set of LRE. We chose to work with the original formulation as it allows for a characterization in case of multiple LRE.

yields

$$m_i > 2kq^* + \frac{2kc}{1+\alpha-\beta} := m^B(c,k).$$

A B-player will remain a B-player if $m_i < m^B(c,k)$, and will randomize between the two strategies if $m_i = m^B(c,k)$. Note that $m^B(c,k)$ is defined as the number of A-players such that a player with less than $m^B(c,k)$ A-neighbors chooses to stay at B with certainty and, thus, cannot exceed $2k$.

Note that $m^A(0,k) = m^B(0,k) = 2kq^*$, i.e. in the absence of switching costs the thresholds are the same as in Ellison's (1993) model. For $c > 0$, we have $m^A(c,k) < m^A(0,k) = m^B(0,k) < m^B(c,k)$. Hence, in the presence of switching costs, it takes more players of the other type to induce a switch than in the absence of switching costs. Further, a B-player will require more A-opponents to switch strategies than an A-player requires to stay at her strategy. Likewise, an A-player will switch to B at a lower number of A-opponents than it takes a B-player to remain at her strategy. Thus, switching costs create regions where players with the same distribution of play in their neighborhood but with a different current strategy may behave differently. This may lead to the emergence of non-monomorphic absorbing states where clusters of players with different strategies coexist. In such states all players want to remain at their current strategies, i.e. $m_i > m^A(c,k)$ for all A-players and $m_j < m^B(c,k)$ for all B-players.

In the following, G denotes the set non-monomorphic absorbing states, i.e.

$$G = \{s \in S | s \neq \overrightarrow{A}, \overrightarrow{B}, m_i > m^A(c,k)\ \forall\ i \text{ with } s_i = A, \text{ and } m_j < m^B(c,k)\ \forall\ j \text{ with } s_j = B\}.$$

and an element of this set is denoted by AB . Further, G_ℓ denotes the set of non-monomorphic absorbing states with ℓ A-players (and $N-\ell$ B-players), i.e.

$$G_\ell = \{s \in G | m = \ell\}.$$

3.1 Two-neighbor interaction

In order to build intuition and to highlight the main mechanisms at work, our analysis starts with an informal discussion of the special case where each agent only interacts with her two most immediate neighbors, i.e. $k = 1$. A comprehensive analysis of the case $k \geq 1$ is provided in Section 3.2.

In a first step, let us consider under which circumstances non-monomorphic states are absorbing. To this end, consider states where clusters of A-players and B-players, each of at least size two, alternate, e.g.

$$\ldots BBAABBBAAAA \ldots$$

Players in the middle of such a cluster only interact with players of their own kind and, hence, will never switch. Thus, let us focus on the boundary between two such strings. Note that whenever $m^A(c, 1) < 1$ holds the boundary A-player will keep her strategy. This translates into $2c > 1 - \alpha - \beta$, which is implied by risk dominance of A. Thus, the boundary A-player will remain. Now consider the B-player. Note that if $m^B(c, 1) \leq 1$ holds, the boundary B-player will switch to A with positive probability. This translates into $c \leq \frac{\alpha+\beta-1}{2}$. Thus, provided switching costs are low, the A-cluster will grow contagiously, even in the absence of mistakes. If this condition is violated, $c > \frac{\alpha+\beta-1}{2}$, the boundary B-player will stay a B-player with certainty. This, in turn, implies that for sufficiently high switching costs non-monomorphic states are absorbing.

Surprisingly, switching costs may not only alter the set of absorbing states but may also change the set of LRE. To see this, first note that one can move among the set of non-monomorphic states via a chain of single mutations. More precisely, it is possible to move from a state in G_ℓ to either a state in G_a or in G_b, with $a < \ell < b$ at the cost of one mutation. While, it is clear that one mutation to A (or B) increases (decreases) the number of A-players by one, this initial mutation might also trigger additional changes.

Further, note that in the presence of non-monomorphic states the transition from one monomorphic state to the other can occur via a series of intermediate steps. Which state will be LRE depends on how difficult it is to move from the two monomorphic states, $\overrightarrow{A}$ and $\overrightarrow{B}$ into the set of non-monomorphic states.[8] First, consider states where there is only one A-player.

$$\ldots BBABB \ldots$$

As the A-player has no A-neighbors she will switch to B with positive probability if $m^A(c,1) \geq 0$ which translates into $c \leq 1-\beta$. In this case lonesome A-players will disappear. However, states with two adjacent A-players are absorbing. Conversely, if $c > 1-\beta$ holds, the A-player will keep her strategy and states with lonesome A-players are absorbing. Likewise, consider the case when there is a lonesome B-player.

$$\ldots AABAA \ldots$$

The B-player has two A-neighbors and will switch strategies with positive probability provided that $m^B(c,1) \leq 2$, which can be rewritten as $c \leq \alpha$. However, whenever $c > \alpha$, a lonesome B-player will remain. Note, by risk dominance of A, $\alpha > 1-\beta$. This implies that whenever lonesome B-players will keep their strategy, lonesome A-player will do the same.

Summarizing, if $c \leq \frac{\alpha+\beta-1}{2}$, only the monomorphic states are absorbing and A can spread out contagiously. Thus, $\overrightarrow{A}$ is unique LRE. If $\frac{\alpha+\beta-1}{2} < c \leq \alpha$ and $c \leq 1-\beta$, non-monomorphic states are absorbing and it is possible to move among the non-monomorphic states and from these states to the monomorphic ones via a single mutation chain. It is further possible to move from the two monomorphic states to the set of non-monomorphic states at the cost of two mutations. Thus, one can exhibit

[8] As the non-monomorphic states can be connected to each other and to the monomorphic states via a chain of single mutations which tree will be of minimum cost will be determined by how difficult it is to escape the monomorphic states. The next section elaborates on this in more detail.

A- and B- trees which are of cost smaller than any AB-tree. Hence, $\overrightarrow{A}$ and $\overrightarrow{B}$ are LRE. If, however, $c > 1 - \beta$, moving from $\overrightarrow{A}$ to the set of non-monomorphic states takes two mutations, whereas escaping $\overrightarrow{B}$ is possible at the cost of one mutation. Thus, in this case one can exhibit A-trees which are of minimum cost, implying that $\overrightarrow{A}$ is unique LRE. If $c > \alpha$, all absorbing states are accessible from each other via a chain of single mutations, implying that all of them are LRE.

Whether it is actually possible that a non-risk dominant convention is LRE does not only depend on the level of switching costs but also on the parameters of the underlying game. To see this point note that both monomorphic states are LRE if $c > \frac{\alpha+\beta-1}{2}$ and $c \leq 1 - \beta$. It, thus, has to be the case that $\frac{\alpha+\beta-1}{2} < 1 - \beta$. This translates into $\alpha + 3\beta < 3$. This condition is fulfilled if the advantage of strategy A over B is not too large, but per se is not related to payoff dominance or risk dominance.[9] Importantly, it may hold if $\alpha > 1$. Thus, even if action A is risk- and payoff- dominant, it might not be unique LRE. We illustrate the set of LRE depending on the level of switching cost in this case in Figure 1. It is interesting to note that the prediction is "non-monotonic" in the level of switching costs. With increasing switching costs the prediction switches from $\overrightarrow{A}$ to $\overrightarrow{A} \cup \overrightarrow{B}$ back to $\overrightarrow{A}$ and finally to $\overrightarrow{A} \cup \overrightarrow{B} \cup G$ in games with $\alpha + 3\beta < 3$.

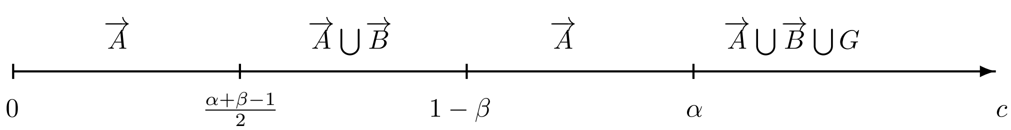

Figure 1: LRE under two player interaction with switching costs and $\alpha + 3\beta < 3$.

3.2 $2k$-neighbor interaction and global interactions

We will now generalize the insights of the two player interaction model to $2k$-neighbor interaction. We show that we can expect similar phenomena as in the simple two-neighbor model for small interaction neighborhoods. However, as the the size of the interaction neighborhood, k, increases switching costs do no longer influence the

[9] If strategy A is sufficiently advantageous compared to B, $\alpha + 3\beta > 3$, it will be uniquely selected up to the point where $c > \alpha$. (where all absorbing states are LRE.)

prediction, with the exception of very high levels of switching costs, where in the absence of noise no player would switch regardless of the distribution of strategies in her neighborhood. The following lemma provides a characterization of the set of absorbing states.

Lemma 1: For positive switching costs,$c > 0$,

i) there are no non-singleton absorbing sets.

ii) the only absorbing states are $\overrightarrow{A}$, $\overrightarrow{B}$ and G.

Proof: To prove the first part consider an absorbing set W. Consider a state $\tilde{s} \in W$ where the number of A-players is maximal. Let $\tilde{m}$ be the number of A-players at this state. It follows that at this state there does not exist a B-player who, when given revision opportunity, switches to A with positive probability. Thus, $m_i < m^B(c,k)$ for all i with $s_i = B$. If it is the case that $m_j > m^A(c,k)$ for all j with $s_j = A$, then $\tilde{s}$ is the only state in W. If $m_j \leq m^A(c,k)$ for some players j with $s_j = A$, we proceed in the following manner. With positive probability, one of these agents receives revision opportunity and switches to B. We reach a new state s'. At this new state there are strictly fewer A-players. Provided that $c > 0$ for the new B-player we have $m_j \leq m^A(c,k) < m^B(c,k)$, implying that she will not switch back. For all old B-players it is still true that $m_i < m^B(c,k)$, implying that none of them will switch. If there is no A-player with $m_j \geq m^A(c,k)$ left, the state s' is absorbing (contradicting that $\tilde{s} \in W$). If there are still such A-players left, we iterate the procedure until we reach an absorbing state, eventually contradicting the assumption $\tilde{s} \in W$.

The second part follows from the definition of $\overrightarrow{A}$, $\overrightarrow{B}$ and G. $\square$

With the help of this lemma we are able to provide the following result.[10]

Proposition 2: In the 2k-neighbor interaction model,

a) if $c \leq \frac{\alpha+\beta-1}{2}$ and $N > k(k+1)$, then $\mathcal{S} = \{\overrightarrow{A}\}$,

[10] In the following we denote by $\lfloor x \rfloor$ the largest integer not greater than x and by $\lceil x \rceil$ the smallest integer not less than x.

b) if $\frac{\alpha+\beta-1}{2} < c \leq \alpha$ and

i) if $\lfloor m^A(c,k) \rfloor = \lfloor 2k - m^B(c,k) \rfloor$, then $\mathcal{S} = \{\overrightarrow{A}, \overrightarrow{B}\}$

ii) if $\lfloor m^A(c,k) \rfloor < \lfloor 2k - m^B(c,k) \rfloor$, then $\mathcal{S} = \{\overrightarrow{A}\}$, and

c) if $c > \alpha$, then $\mathcal{S} = \{\overrightarrow{A}, \overrightarrow{B}\} \cup G$.

Proof: For part a) note if $c \leq \frac{\alpha+\beta-1}{2}$, one has $m^B(c,k) \leq k$, implying that a B-player switches to A with positive probability whenever half (or more) of her $2k$-neighbors choose A. Thus, A may spread contagiously and we are back in the model outlined by Ellison (1993), where $\mathcal{S} = \{\overrightarrow{A}\}$ if $N > k(k+1)$. [11]

We now consider the case where $c > \frac{\alpha+\beta-1}{2}$. Here we have $m^B(c,k) > k$. Thus, B-players will no longer switch if they have half of their neighbors playing A. This implies A can no longer spread out contagiously. Further, non-monomorphic states are now absorbing, meaning that the set G is non-empty.

We next show that it is possible to move from an absorbing state $AB \in G_\ell$ to either a state in G_a or in G_b, with $a < \ell < b$ at the cost of one mutation. We will show that there exists an A- (and a B-player) such that if she mutates to B (to A), she will not switch back and no other player will switch to A (to B). By the definition of G_ℓ we have $m_i > m^A(c,k)$ for all i with $s_i = A$ and $m_j < m^B(c,k)$ for all j with $s_j = B$. Consider now an A-player i whose adjacent neighbor j is playing B. As they are direct neighbors they have only one player who is not a joint neighbor. Call i's disjoint neighbor $\tilde{i}$ and j's disjoint neighbor $\tilde{j}$. Further j also faces i who is an A-player. It follows that j faces either the same number of A-neighbors as i (if $s_{\tilde{i}} = A$ and $s_{\tilde{j}} = B$), has one more A-neighbors than i (if $s_{\tilde{i}} = s_{\tilde{j}}$), or two more A-neighbors (if $s_{\tilde{i}} = B$ and $s_{\tilde{j}} = A$). Thus, $m_j \in \{m_i, m_i+1, m_i+2\}$. Assume that j mutates to A. Since $m_j \geq m_i > m^A(c,k)$ she will not switch back. Further, as there are now more A-players, none of the old A-players will switch, showing that we will reach a state G_b with $b > \ell$. An analogous argument can be used to show that it is

[11] Note that we have a model with positive inertia whereas Ellison's model features strategy adjustment in each round. See Weidenholzer (2010) for a discussion of the model with inertia.

also possible with one mutation to move to a state G_a with $a < \ell$.

Now consider $\overrightarrow{B}$. We want to find the minimum number of mutations required for a transition from $\overrightarrow{B}$ to a state in the set G. Let $C(\overrightarrow{B}, AB)$ denote this number. Recall that $m^A(c,k)$ is defined such that if a player has strictly more than $m^A(c,k)$ A-neighbors, she will strictly prefer to stay at A. If $m^A(c,k) < 0$, we have that an A-player remains even if she does not have an A neighbor. Thus, if $m^A(c,k) < 0$, one mutation is enough to move from $\overrightarrow{B}$ to a state in G_1. Now consider $m^A(c,k) \geq 0$. First, consider the case where $m^A(c,k) \notin \mathbb{Z}$ (where $\mathbb{Z}$ denotes the integers). In this case, if $\lceil m^A(c,k) \rceil + 1$ adjacent players mutate to A each of them will have $\lceil m^A(c,k) \rceil > m^A(c,k)$ players choosing B. Thus, none of them will switch and we have reached an absorbing state in the set $G_{\lceil m^A(c,k) \rceil + 1}$. Note that if less than $\lceil m^A(c,k) \rceil + 1$ players switch to A, all of them will switch back when given revision opportunity. It follows that $C(\overrightarrow{B}, AB) = \max\{\lceil m^A(c,k) \rceil, 0\} + 1$ for $m^A(c,k) \notin \mathbb{Z}$. Now consider $m^A(c,k) \in \mathbb{Z}$. In this case for all A players to stay with probability one each of them needs strictly more than $m^A(c,k)$ A-neighbors. Thus, if $m^A(c,k) + 2$ players switch to A, each of them will have $m^A(c,k) + 1$ neighbors playing A and will not switch back with positive probability. Thus, $C(\overrightarrow{B}, AB) = \max\{m^A(c,k) + 1, 0\} + 1$ for $m^A(c,k) \in \mathbb{Z}$. Summing up, we have

$$C(\overrightarrow{B}, AB) = \begin{cases} \max\{\lceil m^A(c,k) \rceil, 0\} + 1, & \text{if } m^A(c,k) \notin \mathbb{Z} \\ \max\{m^A(c,k) + 1, 0\} + 1, & \text{if } m^A(c,k) \in \mathbb{Z} \end{cases} .$$

This can be written as $C(\overrightarrow{B}, AB) = \max\{\lfloor m^A(c,k) \rfloor + 1, 0\} + 1$.

Conversely, consider the convention $\overrightarrow{A}$. We aim to understand how many mutations to B we need so that the new B-players will keep their strategy with certainty. If $m^B(c,k) > 2k$, this would be the case even if all neighbors choose A. Thus, one mutation is enough to move from $\overrightarrow{A}$ to a state in G_{N-1} whenever $m^B(c,k) > 2k$. Assume $m^B(c,k) \leq 2k$. Now a B-player will keep her strategy whenever $m_i < m^B(c,k)$. Initially the B-players had $2k$ A-neighbors. Thus, each of them needs strictly more than $2k-m^B(c,k)$ of their neighbors to play B to keep their strategy with probability one. Again, let us distinguish the cases $2k-m^B(c,k) \in \mathbb{Z}$ and $2k-m^B(c,k) \notin \mathbb{Z}$. In the latter case with $\lceil 2k-m^B(c,k) \rceil + 1$ mutations one can move from $\overrightarrow{A}$ to a state in the set $G_{N-\lceil 2k-m^B(c,k)\rceil -1}$. Thus, $C(\overrightarrow{A}, AB) = \max\{\lceil 2k-m^B(c,k)\rceil, 0\} + 1$. If $2k-m^B(c,k) \in \mathbb{Z}$, we need $2k-m^B(c,k)+2$ mutations to ensure that each B player has more than $2k-m^B(c,k)$ neighbors playing B. As above, the cases $2k-m^B(c,k) \in \mathbb{Z}$ and $2k-m^B(c,k) \notin \mathbb{Z}$ can be unified by using $C(\overrightarrow{A}, AB) = \max\{\lfloor 2k-m^B(c,k)\rfloor + 1, 0\} + 1$.

Finally, let us determine the set of LRE. Let L denote the number of non-monomorphic absorbing states. Thus, together with the states $\overrightarrow{A}$ and $\overrightarrow{B}$ there are $L+2$ absorbing states. We can connect all L AB states to each other and to $\overrightarrow{A}$ and $\overrightarrow{B}$ via a chain of single mutations. Further, we can move from $\overrightarrow{B}$ into the class of AB states at the cost of $C(\overrightarrow{B}, AB)$. Thus, we can exhibit minimum A-trees of cost $L + C(\overrightarrow{B}, AB)$. Likewise, the minimum B-trees have cost $L + C(\overrightarrow{A}, AB)$. Further, for each state $AB \in G$ we can exhibit a minimum cost tree of cost $L - 1 + C(\overrightarrow{A}, AB) + C(\overrightarrow{B}, AB)$.

First note that if $c > \alpha$, we have $m^A(c,k) < 0$ and $m^B(c,k) > 2k$. It follows $C(\overrightarrow{A}, AB) = C(\overrightarrow{B}, AB) = 1$. Thus, the minimum cost $\overrightarrow{A}$-, the $\overrightarrow{B}$-, and all minimum cost AB-trees have cost $L+1$. Thus, $\mathcal{S} = \{\overrightarrow{A}, \overrightarrow{B}\} \cup G$.

Now, consider $\frac{\alpha+\beta-1}{2} < c \leq \alpha$. Observe that $\lfloor 2k - m^B(c,k)\rfloor = \lfloor 2k(1-2q^*) + m^A(c,k)\rfloor \geq \lfloor m^A(c,k) \rfloor$. Thus, $C(\overrightarrow{A}, AB) \geq C(\overrightarrow{B}, AB)$. So, we either have $C(\overrightarrow{A}, AB) > C(\overrightarrow{B}, AB)$ in which case $\mathcal{S} = \overrightarrow{A}$ or $C(\overrightarrow{A}, AB) = C(\overrightarrow{B}, AB)$ in which case $\mathcal{S} = \overrightarrow{A} \cup \overrightarrow{B}$. $\square$

Thus, the presence of switching costs may imply that under local interactions the risk dominant convention is no longer unique LRE. Let us provide some technical intuition for this result. First, if $c \leq \frac{\alpha+\beta-1}{2}$, the risk dominant strategy may still spread contagiously and, thus, remains unique LRE. For $\frac{\alpha+\beta-1}{2} < c \leq \alpha$ there exist absorbing AB states. Whether the risk dominant or the payoff dominant convention is

LRE boils down to the question from which of the two conventions it is more difficult to move to the set of AB-states. This is measured by the numbers $C(\overrightarrow{A}, AB)$ and $C(\overrightarrow{B}, AB)$ which are in turn rounded down values of the functions $2k - m^B(c,k) + 2$ and $m^A(c,k) + 2$. Risk dominance implies that $C(\overrightarrow{B}, AB) \leq C(\overrightarrow{A}, AB)$. Thus, the risk dominant convention is always contained in the set of LRE. The functions $2k - m^B(c,k) + 2$ and $m^A(c,k) + 2$ only differ by a constant and are linearly decreasing in the switching costs. It may very well be the case that the rounded down values are the same, $C(\overrightarrow{A}, AB) = C(\overrightarrow{B}, AB)$. In this case both conventions turn out to be LRE. Finally, for $c > \alpha$ we have that agents will not switch strategies, no matter what the distribution of strategies among their neighbors is and all absorbing states turn out to be LRE.

In Figure 2, we plot the transition costs from either convention to the set of non-monomorphic states as a function of the switching costs. Whenever $C(\overrightarrow{A}, AB)$ lies above $C(\overrightarrow{B}, AB)$ the convention $\overrightarrow{A}$ is unique LRE. When $C(\overrightarrow{A}, AB)$ and $C(\overrightarrow{B}, AB)$ coincide both conventions, $\overrightarrow{A}$ and $\overrightarrow{B}$, are LRE. When the two functions are equal to one, both conventions, $\overrightarrow{A}$ and $\overrightarrow{B}$, and the set of non-monomorphic states G are LRE. Note that as in the two player interaction case the prediction is non-monotonic in the level of switching costs. In particular, the prediction that the risk dominant convention is unique LRE and the prediction that both of them are LRE alternate k-times.

The following corollary explores the circumstances under which switching costs may influence the set of LRE. In case switching cost may change the prediction, it shows that the prediction will be non-monotonic as switching costs vary.

Corollary 3: If $\frac{\alpha+\beta-1}{2} < c \leq \alpha$ and
a) if $2k(1-2q^*) \geq 1$, then $\mathcal{S} = \overrightarrow{A}$
b) if $2k(1-2q^*) < 1$, then there exist thresholds $\bar{c}^{k+1} < \underline{c}^k < \bar{c}^k < \underline{c}^{k-1} < \bar{c}^{k-1} < \ldots < \underline{c}^1 < \bar{c}^1$ (with $\bar{c}^{k+1} = \frac{\alpha+\beta-1}{2}$ and $\bar{c}^1 = \alpha$) such that if $c \in (\bar{c}^{\ell+1}, \underline{c}^{\ell}]$ for $\ell = 1, 2, \ldots, k$, then $\mathcal{S} = \overrightarrow{A} \cup \overrightarrow{B}$ and $\mathcal{S} = \overrightarrow{A}$ otherwise.

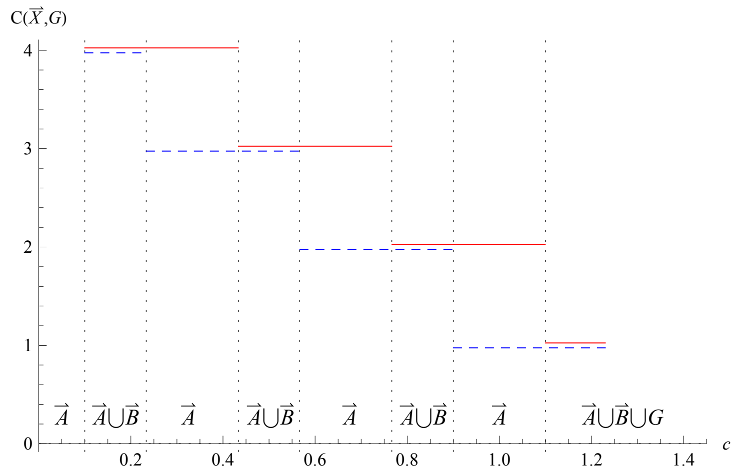

Figure 2: LRE in the game $[\alpha, \beta] = [1.1, 0.1]$ with interaction radius $k = 3$. The solid line plots the transition costs $C(\vec{A}, AB)$ and the dashed line plots the transition costs $C(\vec{B}, AB)$. Whenever $C(\vec{A}, AB)$ lies above $C(\vec{B}, AB)$ the convention $\vec{A}$ is unique LRE. When $C(\vec{A}, AB)$ and $C(\vec{B}, AB)$ coincide both conventions, $\vec{A}$ and $\vec{B}$, are LRE. When the two functions are equal to 1, both conventions, $\vec{A}$ and $\vec{B}$, and the set of non-monomorphic states G are LRE.

Proof: Consider case bii) in the previous Proposition. First, note that $2k - m^B(c,k) = m^A(c,k) + 2k(1-2q^*)$. Thus, the functions $m^A(c,k)$ and $2k - m^B(c,k)$ only differ by the constant $2k(1-2q^*)$. Risk dominance implies $2k(1-2q^*) > 0$. Further, note that $m^A(c,k)$ (and thus also $2k - m^B(c,k)$) is linearly decreasing in c.

Consider part a). Note if $2k(1-2q^*) \geq 1$, then $\lfloor 2k - m^B(c,k) \rfloor > \lfloor m^A(c,k) \rfloor$.

Now consider part b). Let $\underline{c}^\ell$ be the value of switching costs c that solves $m^A(c,k) + 1 = \ell$. Note that $\lfloor m^A(c,k) + 1 \rfloor = \ell$ for $\underline{c}^{\ell+1} < c \leq \underline{c}^\ell$. Likewise, define $\bar{c}^\ell$ to be the value of switching costs c for which $2k - m^B(c,k) + 1 = \ell$. We have $\lfloor 2k - m^B(c,k) + 1 \rfloor = \ell$ for $\bar{c}^{\ell+1} < c \leq \bar{c}^\ell$.

As $2k - m^B(c,k) = m^A(c,k) + 2k(1-2q^*) > m^A(c,k)$ and $m^A(c,k)$ is decreasing in c it follows that $\underline{c}^\ell < \bar{c}^\ell$. Further, note that for $2k(1-q^*) < 1$ one has $m^A(c,k) + 1 < m^A(c,k) + 2k(1-2q^*) + 1 < m^A(c,k) + 2$. Thus, $\underline{c}^\ell < \bar{c}^\ell < \underline{c}^{\ell-1}$. The last two observations imply $\bar{c}^{k+1} < \underline{c}^k < \bar{c}^k < \underline{c}^{k-1} < \bar{c}^{k-1} < \ldots < \underline{c}^1 < \bar{c}^1$. Now note that $\lfloor m^A(c,k) + 1 \rfloor = \lfloor 2k - m^B(c,k) + 1 \rfloor = \ell$ if $c \in (\underline{c}^{\ell+1}, \underline{c}^\ell]$ and $c \in (\bar{c}^{\ell+1}, \bar{c}^\ell]$. This is the case for $c \in (\bar{c}^{\ell+1}, \underline{c}^\ell]$. On the contrary, if $c \in (\underline{c}^\ell, \bar{c}^{\ell-1}]$, then $\lfloor m^A(c,k) + 1 \rfloor = \ell - 1 < \ell = \lfloor 2k - m^B(c,k) + 1 \rfloor$. $\square$

The main idea behind this corollary is that the functions $2k - m^B(c,k) + 2$ and $m^A(c,k) + 2$ only differ by the constant $2k(1 - 2q^*)$. If this constant is greater than or equal to one, we will have $C(\overrightarrow{B}, AB) = \lfloor m^A(c,k) + 2 \rfloor < C(\overrightarrow{A}, AB) = \lfloor 2k - m^B(c,k) + 2 \rfloor$, regardless the level of switching costs. If, however, this constant is smaller than one, there exist levels of switching costs for which $\lfloor 2k - m^B(c,k) \rfloor = \lfloor m^A(c,k) \rfloor$. Further, note that if there exists a range of switching costs for which, e.g., $\lfloor 2k - m^B(c,k) \rfloor = \lfloor m^A(c,k) \rfloor = 1$, then due to the stepwise nature of the floor function, there also exists a range of switching costs for which $\lfloor 2k - m^B(c,k) + 2 \rfloor = \lfloor m^A(c,k) + 2 \rfloor = r$ for every $r \in \mathbb{Z}$. Thus, the prediction that the convention $\overrightarrow{A}$ is unique LRE and the prediction that both conventions, $\overrightarrow{A}$ and $\overrightarrow{B}$, are LRE alternate k times as c increases (and $C(\overrightarrow{B}, AB)$ and $C(\overrightarrow{A}, AB)$ decrease from $k+1$ to 1). Finally, it is interesting to note that since $\underline{c}^1 = 1 - \beta$ and $\bar{c}^1 = \alpha$, for $c \in (1 - \beta, \alpha]$ one has $\mathcal{S} = \overrightarrow{A}$. Further, if $c > \alpha$, one has $\mathcal{S} = \{\overrightarrow{A}, \overrightarrow{B}\} \cup G$. Hence, just before the model's prediction includes all absorbing states it uniquely selects the risk dominant convention.

A straightforward implication of the first part of the corollary is that if agents interact with sufficiently many other agents (k large) or if the risk dominant action has a relatively large basin of attraction (q small), switching costs do not influence the prediction. The second part of the corollary implies that if agents interact only with a few other agents (k small) and/or the risk dominant action's basin of attraction is relatively small (q close to $\frac{1}{2}$), then the prediction may not be unique and moreover is non-monotonic in the level of switching costs.

Finally, note that it is possible to reconcile our findings with the results of Norman (2009) by simply setting $k = \frac{N-1}{2}$, thus, obtaining a model of global interactions. For small populations switching costs may very well have an impact on the set of LRE. However, in large populations, as considered by Norman (2009), the prediction is robust to switching costs. In this case, switching costs speed up convergence but do not alter the long run behavior of the population.

4 Conclusion

We have established that under local interactions the set of LRE may be altered by the presence of switching costs. In particular, the risk dominant convention may no longer be unique LRE. If, however, agents interact with sufficiently many other agents our critique does not apply and risk dominant conventions are still uniquely selected.

One question that immediately comes to mind is whether our findings hold in a more general context. In the context of this paper switching costs played the following role: i) Under switching costs non-monomorphic states may become absorbing. ii) Switching costs may change the transition costs, measured in the number of required mistakes, with which these non-monomorphic states can be accessed from the monomorphic ones. Rounding up, when calculating switching costs, may lead to the effects outlined in this paper. If the number of required mistakes is relatively small, the effect of rounding up will be most pronounced. However, for a large number of required mistakes these effects will be most likely negligible. We, thus, conjecture that switching costs will play a similar role in models where only a relatively small number of mutations is needed to move from one convention to another. There are two natural dimensions along which our results might be generalized. First, we expect switching costs to impact the long run prediction in the circular city model of local interactions if we move beyond the class of 2×2-coordination games (as in e.g Ellison 2000, Alós-Ferrer and Weidenholzer (2005)). Secondly, switching costs will also influence the prediction in models where the way in which agents interact with each other implies that only few mistakes are necessary to move among conventions. Examples include the torus model outlined in Ellison (2000), multiple location models (as in Anwar (2002), Ely (2002), Blume and Temzelides (2003), Shi (2014), Alós-Ferrer and Kirchsteiger (2010)), network formation models under asynchronous adjustments of links and actions (see Jackson and Watts (2002)) or under constrained interactions (as in Staudigl and Weidenholzer (2014)).

Admittedly, the integer problem that is driving our results is an artefact of the uniform noise approach. While other learning models such as the logit dynamics as advocated by e.g. Blume ((1993), (1995)) do not face this problem, their predictions sometimes may depend on other specifics such as the timing of revision opportunities or tie breaking assumptions (see Alós-Ferrer and Netzer (2014) for a discussion).[12]

References

Alós-Ferrer, C. (2004), 'Cournot vs. Walras in Oligopoly Models with Memory', *International Journal of Industrial Organization* 22, 193–217.

Alós-Ferrer, C. & Kirchsteiger, G. (2010), 'General Equilibrium and the Emergence of (Non) Market Clearing Trading Institutions', *Economic Theory* 44, 339–360.

Alós-Ferrer, C. & Netzer, N. (2014), 'Robust stochastic stability', *Economic Theory* pp. 1–27.

Alós-Ferrer, C. & Shi, F. (2012), 'Imitation with asymmetric memory', *Economic Theory* 49(1), 193– 215.

Alós-Ferrer, C. &Weidenholzer, S. (2006), 'Imitation, Local Interactions, and Efficiency', *Economics Letters* 93, 163–168.

Alós-Ferrer, C. & Weidenholzer, S. (2007), 'Partial Bandwagon Effects and Local Interactions', *Games and Economic Behavior* 61, 1–19.

Alós-Ferrer, C. & Weidenholzer, S. (2008), 'Contagion and Efficiency', *Journal of Economic Theory* Volume 143, 251–274.

Anwar, A. W. (2002), 'On the Co-existence of Conventions', *Journal of Economic Theory* 107, 145–155.

Apesteguia, J., Huck, S., Oechssler, J. & Weidenholzer, S. (2010), 'Imitation and the evolution of walrasian behavior: Theoretically fragile but behaviorally robust', *Journal of Economic Theory* 145(5), 1603–1617.

Bergin, J. & Lipman, B. L. (1996), 'Evolution with State-Dependent Mutations', *Econometrica* 64, 943–956.

[12]Maes and Nax (2014) experimentally assess whether uniform- or logit- noise provides a better fit to actual deviation from best response learning. Noise levels turn out to decrease with the payoff loss implied by a deviation, providing support for logit noise. However, for low implied deviation costs (as in the interesting parameter range of our paper) a constant error model may offer a better fit than logit noise.

Blume, A. & Temzelides, T. (2003), 'On the Geography of Conventions', *Economic Theory* 22, 863–873.

Blume, L. (1993), 'The Statistical Mechanics of Strategic Interaction', *Games and Economic Behavior* 5, 387–424.

Blume, L. (1995), 'The Statistical Mechanics of Best-Response Strategy Revision', *Games and Economic Behavior* 11, 111–145.

Ellison, G. (1993), 'Learning, Local Interaction, and Coordination', *Econometrica* 61, 1047–1071.

Ellison, G. (2000), 'Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution', *Review of Economic Studies* 67, 17–45.

Ely, J. C. (2002), 'Local Conventions', *Advances in Theoretical Economics* 2, 1–30.

Eshel, I., Samuelson, L. & Shaked, A. (1998), 'Altruists, Egoists, and Hooligans in a Local Interaction Model', *The American Economic Review* 88, 157–179.

Foster, D. & Young, P. (1990), 'Stochastic evolutionary game dynamics', *Theoretical population biology* 38(2), 219–232.

Freidlin, M. & Wentzell, A. (1988), Random Perturbations of Dynamical Systems, 2nd Ed., Springer Verlag, New York.

Fudenberg, D. & Levine, D. (1998), 'Learning in Games', *European Economic Review* 42, 631–639.

Harsanyi, J. & Selten, R. (1988), A General Theory of Equilibrium Selection in Games, The MIT Press, Cambridge, MA.

Jackson, M. O. & Watts, A. (2002), 'On the formation of interaction networks in social coordination games', *Games and Economic Behavior* 41, 265–291.

Kandori, M., Mailath, G. J. & Rob, R. (1993), 'Learning, Mutation, and Long Run Equilibria in Games', *Econometrica* 61, 29–56.

Kim, C. & Wong, K. (2010), 'Long-run equilibria with dominated strategies', *Games and Economic Behavior* 68, 242–254.

Lee, I. H., Szeidl, A. & Valentinyi, A. (2003), 'Contagion and state dependent mutations', The B.E. *Journal of Theoretical Economics* 0(1).

Maes, M. & Nax, H. H. (2014), 'A behavioral study of'noise'in coordination games', Available at SSRN 2521119 .

Mengel, F. (2009), ‘Conformism and cooperation in a local interaction model’, *Journal of Evolutionary Economics* 19(3), 397–415.

Norman, T. W. (2009), ‘Rapid evolution under inertia’, *Games and Economic Behavior* 66(2), 865–879.

Robson, A. J. & Vega-Redondo, F. (1996), ‘Efficient Equilibrium Selection in Evolutionary Games with Random Matching’, *Journal of Economic Theory* 70, 65–92.

Samuelson, L. (1997), Evolutionary Games and Equilibrium Selection, The MIT Press, Cambridge, Massachusetts.

Shi, F. (2014), ‘Long-run technology choice with endogenous local capacity’, *Economic Theory* pp. 1–23.

Staudigl, M. & Weidenholzer, S. (2014), ‘Constrained interactions and social coordination’, *Journal of Economic Theory* 152, 41–63.

Tanaka, Y. (1999), ‘Long Run Equilibria in an Asymmetric Oligopoly’, *Economic Theory* 14, 705–715.

Vega-Redondo, F. (1997), ‘The Evolution of Walrasian Behavior’, *Econometrica* 65, 375–384.

Weidenholzer, S. (2010), ‘Coordination games and local interactions: A survey of the game theoretic literature’, *Games* 1(4), 551–585.

Weidenholzer, S. (2012), ‘Long-run equilibria, dominated strategies, and local interactions’, *Games and Economic Behavior* 75(2), 1014–1024.

Young, P. (1993), ‘The Evolution of Conventions’, *Econometrica* 61, 57–84.

Schelling's Model Revisited: From Segregation to Integration

Abstract

Schelling (1969, 1971) presents a microeconomic model showing that the individual preferences can drive an integrated city into a rather segregated city, though no player prefers segregation. We study the impact of location heterogeneity on neighborhood segregation in the one-dimensional Schelling residential model. We model location heterogeneity by introducing an advantageous node, in which a player's utility is impartial to the composition of her neighborhood. When every player interacts with two neighbors, we find that one advantageous node in the circular city will lead to a result that segregation is no longer the unique Long-run Equilibria. When players interact with more neighbors, more advantageous nodes are necessary to obtain the same result.

1 Introduction

Residential segregation — the concentration of ethnic, or socioeconomic groups in particular neighborhoods of a city or metropolitan area — is widely perceived as the antithesis of successful integration. This phenomenon is associated with negative outcomes for minorities in terms of academic performance, education attainment, employment and criminal behaviors.[13]

The emergence and persistence of residential segregation may be the result of many factors: economic differences between racial and ethnic groups, housing affordability, different preferences for neighborhood composition, the nature of the urban structure (including job location), and public and private discrimination. Thomas Schelling (1969, 1971a, 1971b, 1978) introduces a model of residential segregation, in which he explores the link between individual preferences and residential segregation. In his model, a player's utility entirely depends on the composition of her neighborhood. Every player is assumed to be satisfied if no more than half of her neighbors are of the opposite type. Unsatisfied players occasionally receive the opportunity to revise their location and move to a satisfactory position. Using an inductive approach, Schelling demonstrates that equilibrium configuration shows high levels of segregation. This result has shown to be robust even if every player prefers integrated neighborhoods (Zhang 2004b; Pancs and Vriend 2007). Young (1998) analytically studies Schelling's model employing the techniques of evolutionary game theory, and proved that the segregated states are the only long-run outcome of a perturbed myopic best-response dynamics.

The present paper aims to understand the role of location heterogeneity in neighborhood segregation. We study a model akin to the one proposed by Young (1998), where players of two types are located in a circular city, each of them interacts with

[13] See Coleman (1966), Bankston and Caldas (1996), Charles and Dinwiddie (2004) and Massey and Fischer (2006) for the influence of residential segregation on academic performance. See Mayer (2002) and Massey and Denton (1993) for the influence of residential segregation on education attainment. See Jencks and Mayer (1990) and Cutler and Glaeser (1997) for the influence of residential segregation on employment. See Shihadeh and Flynn (1996), Collins and Williams (1999) and Krivo and Peterson (1996) for the influence of residential segregation on criminal behaviors. See also Wilson (2012) for other outcomes that result from residential segregation.

two immediate neighbors. We model location heterogeneity by introducing an advantageous node to the city. Player in the advantageous node is satisfied regardless of her neighborhood. The rest players are satisfied if at least one of their neighbors is of own type, and are unsatisfied otherwise. We motivate our study of location heterogeneity by the fact that dwelling units are differentiated in a multitude of dimensions, including their own quality, such as size, age, type, as well as neighborhood effect. An advantageous node represents a dwelling unit which is superior in its own quality, so that people who live in it will ignore the potential disadvantage of neighborhood effect.

The key feature of the advantageous node is that a player residing there is impartial to the composition of her neighborhood. With this advantageous node, one can stay in a neighborhood without neighbors of her own type. This will create an intermediate state, which allow the transitions from segregated states to non-segregated states can be induced via step-by-step evolution as outlined in Ellison (2000). We find that when there is one advantageous node in the city, the long-run prediction of the model will include more integrated neighborhood. Furthermore, when each player interacts with more neighbors on her both sides, the transition from segregated states to non-segregated states requires more mistakes at once, which implies that a cluster of advantageous nodes is necessary to bridge the gap.

The remainder of this paper is structured as follows. The related previous literature will be reviewed in the next section. Section 3 presents the model and discusses the main techniques used. Section 4 presents the main result when players interact with two neighbors. Section 5 extends the model to the case where players interact with more neighbors. Section 6 uses agent-based simulations to assess the quantitative implication of our model. Section 7 concludes.

2 Literature review

Although documented earlier in a literature in sociology (Duncan and Duncan 1957 and Taeuber and Taeuber 1965), it is since Schelling publishes a series of papers (Schelling 1969, 1971a, 1971b, 1972) and the book *Micromotives and Macrobehavior* (1978) that the causes and mechanism of residential segregation have been discussed in an analytical framework. Schelling first presents a one-dimensional model (Schelling (1969)). In this model, players of two types (O and X) are distributed along a linear city, and have the four players on their either side as their neighborhood. Each player can choose his location, and aims to avoid being a minority in his neighborhood. If a player is in the minority, he will insert himself into a satisfied position when given opportunity of revision. Driven by this micromotive, the linear city will unravel into a highly segregated state, even though no individual prefers segregation to integration. In his subsequent works, Schelling considers variations of this model. Schelling (1971a, 1971b) presents a two-dimensional version, where players live on a checkerboard, and some of the cells in the checkerboard are left unoccupied. In this checkerboard model, a player's neighborhood is defined as the eight players around him, which is the so-called Moore neighborhood. An unsatisfied player, who is in the minority in his neighborhood, will move to a vacant cell with satisfactory neighborhood provided it exists, when given opportunity to revise. Schelling (1971b, 1978) also considers different discriminatory preferences, in which the players' tolerance threshold for neighbors of the opposite type increases to 2/3, and the society starts with a perfectly integrated board. With all these variations, the city with two types of players inevitably reaches a highly segregated state. In all these residential segregation models, Schelling focuses on individual preferences, which give impetus to segregation at a global level, even though no individual agent strictly prefers this.

Later the model is tested empirically, mostly by sociologists and geographers. In a seminal contribution in geography, Clark (1991) tests Schelling's model with the data from surveys of residential preferences. He finds support for it, although the preference schedules derived from surveys have a different form from Schelling's assumption. Cutler, Glaeser and Vigdor (1997) examine segregation in American cities from 1890 to 1990. By comparing inter-generational segregation status in different ethnic minorities, the paper finds that decentralized racism is the main contributing factor to segregation since late 20th century. Farley, Fielding and Krysian (1997) examine the preference hypothesis to segregation by using interview data from four US cities, and find that the whites' willingness to move into a neighborhood is inversely related to the density of blacks living there, and the blacks prefer integrated neighborhoods. Thus, the tolerance of blacks for living with white neighbors is crucial to whether integration is likely.

Young (1998) was the first to solve Schelling's model analytically, adopting the techniques developed in evolutionary game theory. In this book, he presented a simple segregation model, where players of two types distribute on a one-dimensional circle, and proved that segregated states are the unique stochastically stable states in the model, even through a segregated neighborhood is not strictly preferred. Zhang (2004b) extends Young's (1998) result to a two-dimensional case, and formulates neighborhood transition as a spatial game played on a lattice graph. In this paper, preferences for neighborhood composition are represented by payoff functions, which peaks at a perfectly integrated neighborhood and is asymmetric—agents prefer being in the majority over being in the minority. He shows that even if everybody prefers balanced neighborhoods, the segregational pattern emerges and persists regardless of the initial state. Zhang (2004a) enriches the two-dimensional model by adding a simple housing market. This model shows that a slight asymmetry in residential preferences between the two groups still induces endogenous segregation.

Schelling's model was also one of the earliest examples of what today would be called an agent-based model. Epstein and Axtell (1996) demonstrate that Schelling's initial result holds under a wide variety of conditions. Bruch and Mare (2006) indicate that very high segregation occurs only when individual behavior is governed by strict thresholds. When the preference function is continuous, segregation is less likely comparing to a step preference function. Benard and Willer (2007) extend Schelling's model to incorporate the wealth and status of agents and desirability and affordability of residences. Given the effect of status-wealth correlation and a housing market, they find that the greater the status-wealth correlation, the more the agents tend to segregate. Laurie and Jaggi (2003) argue that when individuals are able to observe the neighborhood structure of a wider area, integrated neighborhood may become stable. However, Fossett and Warren (2005) argue that Laurie and Jaggi's (2003) result are driven by the assumption of their model. In Laurie and Jaggi's (2003) model, agents move only when they can improve their utility, and will stay in their slot forever when satisfied. This differs significantly from real residential systems which have continuous residential turnover and movement resulting from demographic processes of migration and household lifecycle dynamics. This implausible assumption prevents segregation taking place in the model.

Pancs and Vriend (2007) consider a variety of network structures, (one- and two-dimensions; checkerboard and torus), and find that segregation is the only long-run outcome under all the specifications, even if individuals strictly prefer perfect integration. However, Pancs and Vriend (2007) present the result analytically on a ring only. They argue that the mechanisms of best-response dynamics are different in one- and two-dimensional models. Although they do not extend their analytical result to two-dimensional context, they obtained results in line with the one-dimensional model by using agent-based simulations in a two-dimensional space. O'Sullivan (2009) presents a model with heterogeneity among the agents. He finds that if there are some agents who are indifferent about their neighborhood compositions, segregation in the city will decrease and the overall utility of the agents will increase.

Our research is based on Young's (1998) one-dimensional model, by adopting the same assumption and techniques as Young (1998), we introduce location heterogeneity in the model, and find that segregation is no longer the unique long-run equilibrium in the model, which could provide a better understanding on the mechanism of Schelling's segregation model.

3 The model

3.1 Basic Setup

We start with a one-dimensional residential model following Young (1998). We consider a society of $2n$ players, who lives in a ring network represented by a $2n$-node cycle. Each individual occupies a node, so there is no vacant node in the city. The type (T) of a player is either A or B. We assume that each type has the same population size. We define the *neighborhood* of a player as her k nearest neighbors who live on her left and right sides, i.e. the neighborhood of a player who lives in node i consists of the players at nodes $N(i) =\{i-k, i-k+1, \ldots, i-1, i+1, \ldots, i+k\}$ (modulo $2n$). We further assume that $2n > 2k+1$, so we can rule out the case where some player is counted in the neighborhood of another player multiple times.

We assume that a player's utility depends on the composition of her neighborhood and the node she lives in, she is either *satisfied* or *unsatisfied* with her position. As a variation of the standard Schelling's model, we introduce an *advantageous node* to the city. If a player is in the *advantageous node*, she will be *satisfied* regardless of her neighborhood. When a player lives in a normal node, we assume that her utility depends on the composition of her neighborhood. Following Schelling (1969), we assume a player is *satisfied* if at least half of her neighbors are of the same type, and *unsatisfied* otherwise.

We define a *cluster* as a set of adjacent players of the same type. We denote a cluster by its relative position to the advantageous node. The cluster containing the advantageous node is cluster 1. Cluster $l+1$ is the l^{th} cluster to the right of cluster 1. Let z denote a state of the city, which specifies how the $2n$ players are arranged in the circular city. We denote by Z the set of the states. We denote by m_l the size of cluster l. We denote by T^* the player of type T, who is located in the advantageous node. Consider a state $z \in Z$, in which the players are arranged as:

$$z = \underbrace{A \vdots A^* \dots A}_{m_1} \underbrace{B \dots B}_{m_2} \underbrace{A \dots A}_{m_3} \underbrace{B \dots B}_{m_4} \dots \underbrace{B \dots B}_{m_M} \underbrace{A \vdots A^* \dots A}_{m_1},$$

where the vertical dashed line indicates the position of the advantageous node, where the city joins. We denote by $m(z) = (m_1, m_2, \dots m_{M-1}, m_M)$, the distribution of the players in state z, where we denote by M the number of clusters in the city.

Different from the original Schelling's model, we have an advantageous node in the city. Every node can be defined according to its relative position to this absolute node. Consider two states z and z', every player in z and z' has the same relative position to other players, but has a different relative position to the advantageous node. In Schelling's original model, where the position of a player is only defined by her relative position to her neighbors, z and z' are exactly the same states. However, these two states are different states in our model.

3.2 The dynamic process

At each round, one pair of players is chosen uniformly at random and exchanges their location if the swap can increase the aggregate utility of the chosen players. Apparently, there can be positive gain from the swap only if the two chosen players are of different types. If two players of the same type exchange, both of them will not change their utility before and after the swap, and there will be no gain from the swap. There are two possible cases when the swap takes place with positive gain: i) one player is unsatisfied and the other is satisfied before the swap, and both are satisfied afterward, or ii) both players are unsatisfied before the swap and both are satisfied afterward. Since we assume that a player is either satisfied or unsatisfied with her position, a swap with positive gain is always Pareto-improving.

3.3 Review of techniques

The city evolves by making swaps with positive gain. In addition, we assume that with fixed probability $\epsilon > 0$, independent across players and across time, a pair of players ignores their prescription and makes a swap with no positive gain in aggregate utility, i.e. they make a mistake or mutate.

Definition 1: A state z is *absorbing* if no alternative state z' can be reached from z without mutations. We denote by Z the set of absorbing states.

The process with mistakes is called *perturbed process*. Under mistakes, the process is irreducible for any two states can be reached from each other. We denote the unique invariant distribution of this process by $\mu(\epsilon)$. As the rate of experimentation converges to 0, the limit invariant distribution $\lim_{\varepsilon\to 0}\mu(\epsilon) = \mu^*$ predicts the process in the long run. The states in the support of μ^* are called *Long Run Equilibria* or stochastically stable states. We use a characterization of the set of LRE due to Freidlin and Wentzell (1988). For each absorbing state z, a z-tree is a set of directed edges such that, from every absorbing state different from z, there is unique directed path in the tree to z. For each edge, the minimal number of mutations that needed for evolving from one absorbing state to another is called transition cost. The resistance of the z-tree is the sum of the transition cost on the edges that compose it. When the probability of error converges to 0, the perturbed process is most likely to follow the paths that lead to the states with least resistance.

Lemma 1: (Freidlin and Wentzell 1988): The LRE of the process is the set of states which have z-trees with the least resistance.

We will also make use of the concept of a mutation-connected component, which will simplify the analysis of the where a class of absorbing states can be reached from each other via a series of single mutations.

Definition 2: A set of absorbing states Z^0 is a *mutation-connected component* if for any $z, z' \in Z^0$, it is possible to go from z to z' through a sequence of single mutation transitions.

Lemma 2: (Nöldeke and Samuelson 1993, Proposition 1): If one state in the mutation-connected component is LRE, so are the states in the same mutation-connected component.

4 Location heterogeneity in the two-neighbor residential model

In this part, we will characterize all the absorbing states in a two-neighbor model with one advantageous node, and identify the LRE of the model. In the next section, we will generalize the model to the case where $k > 1$.

First, we characterize the absorbing states in the model. In our model, a player is satisfied in two cases: i) she lives in a normal node and one of her neighbors is of her own type; ii) she lives in the advantageous node. When a player in a normal node has no neighbor of her own type, she is unsatisfied. We refer to such player as *the isolated player*.

Lemma 3: In the two-neighbor residential model with one advantageous node, the absorbing states of the unperturbed process are the states where there is no isolated player in the normal node. So, the set of absorbing states (Z^0) is:

$$Z^0 = \{z | m_1(z) \geqslant 1,\, m_i(z) \geqslant 2, i = 2, 3, 4 \ldots M\}, \text{ where } M \in \{2, \ldots, 2\left\lfloor \tfrac{n}{2} \right\rfloor\}.$$

Proof: i) Consider a state $z \in Z^0$, in which no player in the normal node is isolated. In this case, every player in the normal nodes has at least one neighbor of the same type, and the player in the advantageous node is satisfied anyway. Thus, every player in the city is satisfied and the state is absorbing.

ii) Consider a state $z' \notin Z^0$, where there exists at least one isolated player in a normal node. Without loss of generality, we assume the isolated player is an A-player,

$$z' = \underbrace{A \ldots A \vdots A^* A \ldots A}_{m_1} \ldots \underbrace{B \ldots B}_{m_i} \underbrace{\mathbf{A}}_{1} \underbrace{B \ldots \mathbf{B}}_{m_{i+2}} \underbrace{A \ldots A}_{m_{i+3}} \ldots \underbrace{B \ldots B}_{m_M} \underbrace{A \ldots A \vdots A^* A \ldots A}_{m_1}.$$

When she is selected, she can exchange her location with a B-player on the border of a B-cluster. These two players are indicated in bold. Since both players involved are satisfied after the swap, they will change the location with probability one. The city will reach a state z'' where no player in the normal nodes is isolated.

$$z'' = \underbrace{A \dots A \vdots A^* A \dots A}_{m_1} \dots \underbrace{B \dots \mathbf{B} \dots B}_{m_i + m_{i+2}} \underbrace{\mathbf{A} A \dots A}_{m_{i+3}+1} \dots \underbrace{B \dots B}_{m_M} \underbrace{A \dots A \vdots A^* A \dots A}_{m_1}.$$

Thus, for any state, when there exists at least one isolated player in the normal node, the state is not absorbing, and will reach an absorbing state with positive probability. □

Now, we categorize the set of absorbing states into subsets based on the number of clusters in the state. We denote by $Z^0(M)$ the set of absorbing states with M clusters. Next, we give the definition of the segregated states.

Definition 3: *A segregated* state, is a state in which there are only two clusters in the city, one of which consists of all the A-players and the other one consists of all the B-players, i.e. $M = 2$.

Thus, the set of segregated states is given by $Z^0(2)$. Correspondingly, in a state, if there are more than two clusters in the city, we call the state a *non-segregated* state.

Before we move to the main result, we present a series of lemmas that show how the city can move among various absorbing states.

In the following lemma, we consider the transitions among the absorbing states with the same number of clusters:

Lemma 4: When $M < n$, the absorbing states in $Z^0(M)$ compose a mutation-connected component.

Proof: Consider an absorbing state $z \in Z^0(M)$, where $M < n$. Without loss of generality, we assume that an A-player occupies the advantageous node, so we have:

$$z = \underbrace{A \ldots A \vdots A^{*} A \ldots A}_{m_1} \underbrace{B \ldots B}_{m_2} \ldots \underbrace{B \ldots B}_{m_M} \underbrace{A \ldots A \vdots A^{*} A \ldots A}_{m_1}.$$

Since $M < n$, there exist one B-cluster and one A-cluster, each of them has more than two players. Here we assume that $m_i^A \geqslant 3$, where $i \in \{1, 3, \ldots, M-1\}$. Consider two adjacent A-clusters, cluster i and cluster $(i+2)$. The city can evolve from an absorbing state z, where $m(z) = (m_1, m_2, \ldots, m_i, m_{i+1}, m_{i+2}, \ldots, m_{M-1}, m_M)$,

$$z = \underbrace{A \ldots A \vdots A^{*} A \ldots A}_{m_1} \underbrace{B \ldots B}_{m_2} \ldots \underbrace{A \ldots A\mathbf{A}}_{m_i} \underbrace{B \ldots B\mathbf{B}}_{m_{i+1}} \underbrace{A \ldots A}_{m_{i+2}} \ldots \underbrace{A \ldots A \vdots A^{*} A \ldots A}_{m_1}$$

to another absorbing state z', where $m(z') = (m_1, m_2, \ldots, (m_i - 1), m_{i+1}, (m_{i+2} + 1), \ldots, m_{M-1}, m_M)$,

$$z' = \underbrace{A \ldots A \vdots A^{*} A \ldots A}_{m_1} \underbrace{B \ldots B}_{m_2} \ldots \underbrace{A \ldots A}_{m_i - 1} \underbrace{\mathbf{B}B \ldots B}_{m_{i+1}} \underbrace{\mathbf{A}A \ldots A}_{m_{i+2}+1} \ldots \underbrace{A \ldots A \vdots A^{*} A \ldots A}_{m_1}$$

with one mutation. By swaps like this, we can rearrange the distribution of A-players among any two adjacent A-clusters. Iterating these single mutations, we can change the distribution of all the A-players among these $M/2$ A-clusters, without affecting the distribution of B-players. Due to symmetry of the types, we can also change the distribution of all the B-players among these $M/2$ B-clusters.

Thus, when $M < n$, all the absorbing states with M cluster are connected via single mutations. $\square$

By Lemma 4, when $M < n$, every absorbing state with the same number of clusters can reach each other via single mutations. This implies that, if an absorbing state $z \in Z^0(M)$ can reach an absorbing state z' via single mutations, every absorbing state in $Z^0(M)$ can reach z' via single mutations; if an absorbing state z' can reach an absorbing state $z \in Z^0(M)$ via single mutations, z' can reach every absorbing state in $Z^0(M)$ via single mutations.

In the following lemma, we consider the transition from the absorbing states with more clusters to an absorbing state with fewer clusters:

Lemma 5: It is possible to move from an absorbing state in $Z^0(M)$ to an absorbing state in $Z^0(M-2)$ via a series of single mutations.

Proof: Consider an absorbing state $z \in Z^0(M)$

$$z = \underbrace{A\dots A\vdots A^* A\dots A}_{m_1}\underbrace{B\dots B}_{m_2}\dots\underbrace{B\dots B}_{m_{i-1}}\underbrace{A\mathbf{A}}_{2}\underbrace{B\dots B}_{m_{i+1}}\dots\underbrace{B\dots B\mathbf{B}}_{m_j}\underbrace{A\dots A}_{m_{j+1}}\dots\underbrace{A\dots A\vdots A^* A\dots A}_{m_1},$$

in which one of the clusters in normal nodes has only two players. Without loss of generality, let $m_i = 2$. When one of the two players in cluster i and a B-player on the border of a B-cluster are selected and swap by mistake, we will reach a state z',

$$z' = \underbrace{A\dots A\vdots A^* A\dots A}_{m_1}\underbrace{B\dots B}_{m_2}\dots\underbrace{B\dots B}_{m_{i-1}}\underbrace{A}_{1}\underbrace{\mathbf{B}B\dots B}_{m_{i+1}+1}\dots\underbrace{B\dots B}_{m_j-1}\underbrace{\mathbf{A}A\dots A}_{m_{j+1}+1}\dots\underbrace{A\dots A\vdots A^* A\dots A}_{m_1},$$

in which cluster i has only one A-player, who is isolated and unsatisfied. The state z' now lies in the basin of attraction of another absorbing state. When given opportunity, the isolated A-player could be selected and swap with another B-player with positive probability and reach an absorbing state $z'' \in Z^0(M-2)$

$$z'' = \underbrace{A\dots A\vdots A^* A\dots A}_{m_1}\underbrace{B\dots B}_{m_2}\dots\underbrace{B\dots B}_{m_{i-1}+m_{i+1}+1}\dots\underbrace{A\dots A\vdots A^* A\dots A}_{m_1}.$$

After these two swaps, cluster i has been eliminated, cluster $(i-1)$ and cluster $(i+1)$ merge into one B-cluster. Therefore, it is possible to move from an absorbing state in $Z^0(M)$ to an absorbing state in $Z^0(M-2)$ via a series of single mutations. $\square$

In the following lemma, we consider the transition from the absorbing states with fewer clusters to an absorbing state with more clusters with the help of the advantageous node.

Lemma 6: When $M \leqslant n-3$, it is possible to move from the absorbing states in

$Z^0(M)$ to the absorbing states in $Z^0(M+2)$ via a series of single mutations.

Proof: When $M = n-3$, in an absorbing state with M clusters, the largest possible cluster can have five players. We consider the set of absorbing states $Z^0(M)$, where $M \leqslant n-3$. In $Z^0(M)$, there exists an absorbing state z, such that (1) the advantageous node is occupied by an A-player, (2) at least two A-neighbors on the left hand side of the advantageous node and (3) at least two A-neighbors on the right hand side of the advantageous node. It implies that cluster 1 has at least five players. The state z is given by

$$z = \underbrace{A\ldots A\vdots \mathbf{A}^* A\ldots A}_{m_1}\underbrace{B\ldots B}_{m_2}\cdots\underbrace{B\ldots B\mathbf{B}}_{m_i}\underbrace{A\ldots A}_{m_{i+1}}\cdots\underbrace{B\ldots B}_{m_M}\underbrace{A\ldots A\vdots A^* A\ldots A}_{m_1}.$$

Since $M \leqslant n-3$, there exists a B-cluster i in z, such that $m_i \geqslant 3$. In the first step, the A^*-player and a B-player on the border of cluster i are selected and swap by mistake, a new absorbing state z' is obtained,

$$z' = \underbrace{A\ldots A\vdots \mathbf{B}^* A\ldots A}_{m_1}\underbrace{B\ldots B}_{m_2}\cdots\underbrace{B\ldots B}_{m_i-1}\underbrace{\mathbf{A}A\ldots A}_{m_{i+1}+1}\cdots\underbrace{B\ldots B}_{m_M}\underbrace{A\ldots A\vdots A^* A\ldots A}_{m_1}.$$

In the second step, an A-player next to the B^*-player and a B-player on the border of the adjacent B-cluster could be selected and swap by mistake. We will reach an absorbing state with $(M+2)$ clusters.

$$z'' = \underbrace{A\ldots A}_{m_t}\vdots\underbrace{B^*\mathbf{B}}_{2}\underbrace{A\ldots \mathbf{A}}_{m_1-m_t-1}\underbrace{B\ldots B}_{m_2-1}\cdots\underbrace{B\ldots B}_{m_M}\underbrace{A\ldots A}_{m_t}\vdots\underbrace{B^*B}_{2}.$$

After these two swaps, we reach a new absorbing state $z'' \in Z^0(M+2)$, where a new cluster 1 with two B-players is created and the original cluster 1 is separated into two parts: cluster 2 and cluster $M+2$. Therefore, when $M \leqslant n-3$, the city can evolve from absorbing states in $Z^0(M)$ to absorbing states in $Z^0(M+2)$ via a series of single mutations. $\square$

Note that in the absorbing state z' in the proof of Lemma 6, the B^*-player has no neighbor of her own type, she is satisfied due to the presence of the advantageous node. If a state is not absorbing when the nodes in the city are homogeneous, and become absorbing only after the advantageous node is introduced, we refer to such state as *the intermediate absorbing state*. With the help of the advantageous node, the intermediate absorbing states are created, so that the transitions from the absorbing states with fewer clusters to an absorbing state with more clusters can be induced via single mutations.

In the following proposition, we will identify the LRE of the model.

Proposition 1: When n is odd, the set of LRE is Z^0, when n is even, the set of LRE is $Z^0/Z^0(n)$.[14]

Proof: i) When i is even, there can be at most n clusters in an absorbing state, the set of absorbing states with n clusters is $Z^0(n)$. We know by Lemma 6 that, when $M \leqslant n-3$, the transition from the absorbing states with fewer clusters to an absorbing state with more clusters be induced via single mutations. However, in an absorbing state in $Z^0(n-2)$, the largest possible cluster can have four players, it is not possible to reach the absorbing states in $Z^0(n)$ via single mutations. By Lemma 5 and Lemma 6, the absorbing states in $Z^0/Z^0(n)$ can reach each other via single mutations, but cannot reach the absorbing states in $Z^0(n)$ via single mutations. So we find that: i) for every absorbing state in $Z^0(n)$, every z-tree has at least one edge with a resistance of two; ii) for every absorbing state in $Z^0/Z^0(n)$, there exists a z-tree in which every edge has a resistance exactly equal to one. Let L denote the total number of absorbing states in the model. All the absorbing states in $Z^0/Z^0(n)$ can form a mutation-connected component, in which every state has a z-tree with the least resistance of $L-1$, the absorbing states in $Z^0(n)$ have z-tree with a resistance no less than L. By Lemma 1 and Lemma 2, the set of LRE is $Z^0/Z^0(n)$.

[14] We denote by A/B the set of all elements that are members of A but not members of B.

ii) When n is odd, there can be at most $(n-1)$ clusters in an absorbing state. A state in $Z^0(n-3)$ can have a cluster with five players, by Lemma 6, We know that the absorbing states in $Z^0(n-3)$ can reach a state in $Z^0(n-1)$ via single mutations. In this case, we know by Lemma 5 and Lemma 6 that all the absorbing states in Z^0 form a mutation-connected component. Thus, all the absorbing states can be connected with single mutations. By Lemma 2, we can conclude that when n is odd, the set of LRE is Z^0. $\square$

In Lemma 4, 5 and 6, we explore how the city evolves among the absorbing states via single mutations. By Lemma 4, we know that absorbing states with the same number of clusters can reach each other via single mutations. By Lemma 5, we can reach from the absorbing states with more clusters to the absorbing states with fewer clusters via single mutations. Without the advantageous node, the transition from the absorbing states with fewer clusters to the absorbing states with more clusters requires at least two mistakes at once. This makes the absorbing states with fewer clusters are easier to reach than to leave. Thus, in Schelling's original model, the absorbing states with fewest clusters — the segregated states — will be the unique LRE. However, once the advantageous node is introduced to the city, the intermediate states are created. By Lemma 6, an absorbing state with fewer clusters can reach an intermediate state, then to an absorbing state with more clusters via single mutations. By Lemma 4, 5, and 6, we find that given the number of clusters is not too large, the states with different number of clusters can reach each other via single mutations. The segregated states are not the unique LRE in the model with advantageous node. In Proposition 1, we find that when n is even, the set of LRE includes all the absorbing state in $Z^0/Z^0(n)$. The states in $Z^0(n)$ have too many clusters, so that transition from $Z^0(n-2)$ to $Z^0(n)$ cannot be induced via single mutations by Lemma 6. When n is odd, all the absorbing states are included in the set of LRE.

5 Location heterogeneity in the $2k$-neighbor residential model

5.1 The $2k$-neighbor model without advantageous node

Next we consider the $2k$-neighbor model. First we assume that $2n > 2k + 1$, for any $k > 1$ to rule out the case where some player is counted in the neighborhood of another player multiple times. In the two-neighbor model, we assume that a player will be satisfied if she is in the majority of her neighborhood, or she is located in an advantageous node. In the $2k$-neighbor case, a player in normal node interacts with $2k$ players on her both sides, she will be satisfied if at least k of them are of her own type, and will be unsatisfied otherwise.

In the rest of the section, we first consider the case where all nodes in the circular city are normal nodes, we identify the absorbing states and LRE in this case. Then, we introduce the advantageous nodes to the city, and discuss the implication of the heterogeneity among locations.

The following lemma characterizes the set of absorbing states in the $2k$-neighbor model when the nodes in the city are homogeneous.

Lemma 7: When the nodes in the city are homogenous, the absorbing states of the $2k$-neighbor model can be categorized into the following two types:
(1) The states in which every cluster has no fewer than $(k + 1)$ players;
(2) The states in which every cluster has m players, where $m < k + 1$, and $k = 2pm$ or $k = (2p + 1)m - 1$, $p = 1, 2, 3 \ldots$, for any $k < \frac{2n-1}{2}$.

Proof: For type (1), consider a cluster with no fewer than $(k + 1)$ players of type $T \in \{A, B\}$. Each player in the cluster has at least k neighbors of the same type within the cluster, so every player in the cluster is satisfied. If every cluster in a state has no fewer than $(k+1)$ players, all players are satisfied, and the state is absorbing.

For type (2), when each cluster has exactly m players, and $k = 2pm$ or $k = (2p + 1)m - 1$, each player has exactly k neighbors of own type, so the states are absorbing.

Next, we show that if in a state, not all of the clusters have fewer than $(k+1)$ players, the state is not absorbing, and will reach an absorbing state with positive probability. Consider a state z, such that at least one of the clusters in z has fewer than $(k+1)$ players, and at least one of the clusters in z has no fewer than $(k+1)$ players. There exist two adjacent clusters in z, one has fewer than $(k+1)$ players, and one has no fewer than $(k+1)$ players. Without loss of generality, we assume the A-cluster (cluster i) has fewer than $(k+1)$ players, and its adjacent B-cluster (cluster $(i-1)$) has no fewer than $(k+1)$ players, i.e. $m_{i-1} \geqslant k+1$, $m_i < k+1$

$$z = \underbrace{A\dots A\vdots A^*A\dots A}_{m_1}\underbrace{B\dots B}_{m_2}\dots\underbrace{B\dots B}_{m_{i-1}}\underbrace{\mathbf{A}A\dots A}_{m_i}\dots\underbrace{B\dots B}_{m_M}\underbrace{A\dots A\vdots A^*A\dots A}_{m_1}.$$

Consider the A-player (indicated in bold) on the border of the cluster i. She has k B-players on the left hand side of her neighborhood, and she will be satisfied only if rest of her neighbors are A-players. Since this cluster i has fewer than $(k+1)$ A-players, this emboldened A-player is unsatisfied, and will change her location with positive probability. After the emboldened A-player change her location with other B-player, the rest A-player in cluster i will be unsatisfied, and will also change their locations successively. Iterating this process, every cluster which has fewer than $(k+1)$ player will be eliminated. Thus, the state z will reach an absorbing state of type (1) with positive probability.

Then, we show that if in a state, every cluster has fewer than $(k+1)$ players, but not every cluster has the same number of players, the state is not absorbing, and will reach an absorbing state with positive probability. We consider the following three cases. Consider a state z', in which every cluster has fewer than $(k+1)$ players. Let cluster ℓ be the largest cluster in z, and there exists at least one cluster has strictly fewer players than cluster ℓ.

i) First, we show that, if every cluster in a state has fewer than $(k+1)$ players, it is absorbing only if every player on the border of a cluster has exactly k neighbors of own type. Consider two adjacent clusters, one is an A-cluster, the other is a B-cluster. Without loss of generality, we assume that the A-player next to the B-cluster – we call her player i – has no fewer than $(k+1)$ A-neighbors. The B-player next to player i – we call her player j – has $(2k-2)$ joint neighbors with player i. Each of the two players has an disjoint neighbor. Since player i has no fewer than $(k+1)$ A-neighbors, there are at least $(k+1)$ A-players among the $(2k-2)$ joint neighbors and player i's disjoint neighbor. Thus, there are at least k A-players in the $(2k-2)$ joint neighbors. Now we consider the player j's neighborhood. There are at least k A-players in the joint neighbors, and player i is also an A-player in player j's neighborhood. In this case, player j has at least $(k+1)$ A-players in her neighborhood. Player j is unsatisfied, and will change her location with positive probability. Once this B-player change her location, the rest B-players in the B-cluster will also change their location successively. The state will reach an absorbing state with positive probability. From this, we know that if every cluster in a state has fewer than $(k+1)$ players, and one of the player on the border of a cluster has no fewer than $(k+1)$ neighbors of own type, the state is not absorbing.

ii) Next, we show that, if every player on the border of a cluster has exactly k neighbors of her own type, and one cluster is strictly larger than some other cluster, the state is not absorbing. We know that every player on the border of a cluster has exactly k neighbors of her own type, and the neighborhood size of every player is the same. Since cluster ℓ is the largest cluster, players in cluster $(\ell+1)$ and cluster $(\ell-1)$ have fewer than k neighbors of their own type. These players will change their locations when given revision opportunity. We will obtain an absorbing state with positive probability.

Thus, when every cluster has fewer than $(k+1)$ players, but not every cluster has the same number of players, the state is not absorbing.

To sum up, the states of type (1) and type (2) are the absorbing states. A state different from type (1) and type (2) are not absorbing, and will evolve to an absorbing state with positive probability.□

Next we identify the LRE of the $2k$-neighbor model. We denote by Z^{k+1} the set of absorbing states. We denote by Z_0^{k+1} the set of absorbing states of type (1), by Z_0^m the set of absorbing states of type (2). As we did in the two-neighbor model, we categorize the absorbing states in Z_0^{k+1} according to the number of clusters in the state. We denote by $Z_0^{k+1}(M)$ the subset of Z_0^{k+1} in which there are M clusters in the city. So the set of segregated states is $Z_0^{k+1}(2)$. Since the transitions depicted in Lemma 4 and Lemma 5 are irrelevant to the advantageous node, we can expect the same transitions in the $2k$-neighbor model without advantageous node. As the immediate extensions of Lemma 4 and Lemma 5, we have the following lemmas.

Lemma 8: In the $2k$-neighbor model, when the nodes in the city are homogenous, the absorbing states in $Z_0^{k+1}(M)$ compose a mutation-connected component, where $M < 2\left\lfloor \frac{n}{k+1} \right\rfloor$.

Lemma 9: In the $2k$-neighbor model, when the nodes in the city are homogenous, it is possible to move from an absorbing state in $Z_0^{k+1}(M)$ to an absorbing state in $Z_0^{k+1}(M-2)$ via a series of single mutations, where $M < 2\left\lfloor \frac{n}{k+1} \right\rfloor$.

Now we identify the LRE of the $2k$-neighbor model when the nodes in the city are homogenous:

Proposition 2: When the nodes in the city are homogenous, the set of LRE in the $2k$-neighbor residential model is the set of segregated states.

Proof: For any absorbing state in Z_0^m, every player has exactly k neighbors of own type. If any two players of different types change their locations, it will make some other players unsatisfied. We consider the following swap: an A-player on the border of an A-cluster and a B-player on the border of a B-cluster are selected and change their location by mistake. Once this swap takes place, the rest A-players in the A-cluster and the rest B-players in the B-cluster have fewer than k neighbors of their own types. These players become unsatisfied and will change their locations with positive probability. This will make more players unsatisfied and change their locations with positive probability, and finally, we will reach a state in Z_0^{k+1}. Thus, to leave an absorbing state in Z_0^m, we need just one mutation. On the other hand, in an absorbing state in Z_0^m, every cluster has strictly fewer than $(k+1)$ players, the absorbing states in Z_0^m have more clusters than the absorbing states in Z_0^{k+1}. To reach an absorbing state in Z_0^m from states in Z_0^{k+1}, we need to increase the number of clusters in the city, which requires strictly more than one player to move at once. Thus, the transition cost from states in Z_0^{k+1} to an absorbing state in Z_0^m is strictly larger than one.

We know by lemma 8 that the states in $Z_0^{k+1}(2)$ compose a mutation-connected component. To leave $Z_0^{k+1}(2)$, we need $(k+1)$ pairs of players change their location at once. Thus, the transition cost from a segregated state to a non-segregated absorbing state is at least $(k+1)$. By Lemma 9, the transition from any non-segregated absorbing state to a segregated state can be induced via a series of single mutations. Let L denote the total number of absorbing state in the model, all the segregated states have a z-tree with least resistance of $L-1$. While for any non-segregated absorbing state, every z-tree has at least one edge which is larger or equal to $(k+1)$, so the resistances of these z-trees are at least $L+k$. Therefore, the set of segregated states ($Z_0^{k+1}(2)$) is the set of LRE. $\square$

As shown in the original Schelling model, when players interact with more neighbors, the set of segregated states is still the unique LRE.

5.2 Multiple advantageous nodes in the $2k$-neighbor model

Recall that in the two-neighbor model, an intermediate state is created with the help of the advantageous node, so that the absorbing states with different number of clusters can be connected via single mutations. It turns out that almost all the absorbing states are included in the set of LRE. However, one advantageous node creates only one intermediate absorbing state. In the $2k$-neighbor model, the transition cost from $Z_0^{k+1}(M)$ to $Z_0^{k+1}(M+2)$ is $(k+1)$. With just one advantageous node, the absorbing states in $Z_0^{k+1}(M)$ cannot reach the states in $Z_0^{k+1}(M+2)$ via single mutations. Therefore, we need k adjacent advantageous nodes to enlarge the set of LRE.

Lemma 10: In the $2k$-neighbor model with k adjacent advantageous nodes, the absorbing states can be categorized into the following three types:

($1'$) The states in which every cluster has no fewer than $(k+1)$ players;

($2'$) The states in which every cluster has m players, where $m < k+1$, and $k = 2pm$ or $k = (2p+1)m - 1$, $p = 1, 2, 3 \ldots$, for any $k < \frac{2n-1}{2}$;

($3'$) The states in which at least one player in the advantageous nodes has fewer than k neighbors of own type, and every cluster containing normal nodes has no fewer than $(k+1)$ players.

Proof: From Lemma 7, we know that the states of type ($1'$) and type ($2'$) are absorbing in the $2k$-neighbor model without advantageous nodes. Since the presence of the advantageous nodes does not change the utility of players in these states, these states are still absorbing in the $2k$-neighbor model with k adjacent advantageous nodes. For the states of type ($3'$), they are not absorbing in the $2k$-neighbor model without advantageous nodes. They become absorbing due to the presence of the advantageous nodes. Consider a state where every cluster containing normal nodes has no fewer than $(k+1)$ players. If in such a state every player in the advantageous node has at least k neighbors of her own type, the state belongs to type ($1'$). If at least one player in the advantageous node has fewer than k neighbors of her own type, she is satisfied due to the existence of the advantageous nodes. Then, the state is absorbing and belongs to type ($3'$). The states different from type ($1'$), type ($2'$) and type ($3'$) are not absorbing, and will evolve to an absorbing state of type ($1'$) or type ($3'$) with positive probability. □

We denote by Z_k^{k+1} the set of absorbing states of type ($1'$), by Z_k^m the set of absorbing states of type ($2'$), by $Z_k^{(k+1)'}$ the set of absorbing states of type ($3'$). Since the states in $Z_k^{(k+1)'}$ is not absorbing when the nodes in the city are homogeneous, and become absorbing after the advantageous nodes are introduced, these states are the intermediate absorbing states. Furthermore, the absorbing states in Z_k^{k+1} can be categorized into subsets based on the number of clusters. We denote by $Z_k^{k+1}(M)$ the set of absorbing states in Z_k^{k+1} with M clusters. Since the transition within $Z_k^{k+1}(M)$ and the transition from $Z_k^{k+1}(M)$ to $Z_k^{k+1}(M-2)$ can be induced via single mutations without the help of the advantageous nodes, Lemma 8 and Lemma 9 can be applied to the $2k$-neighbor model with k adjacent advantageous nodes. So we can have the following lemma immediately:

Lemma 11: In the $2k$-neighbor model with k adjacent advantageous nodes, the absorbing states in $Z_k^{k+1}(M)$ compose a mutation-connected component when $M < 2\left\lfloor \frac{n}{k+1} \right\rfloor$. It is possible to move from an absorbing state in $Z_k^{k+1}(M)$ clusters to an absorbing state in $Z_k^{k+1}(M-2)$ clusters via a series of single mutations, where $M < 2\left\lfloor \frac{n}{k+1} \right\rfloor$.

In the following lemma, we consider the transition from the absorbing states in Z_k^{k+1} with fewer clusters to an absorbing state in Z_k^{k+1} with more clusters.

Lemma 12: In the $2k$-neighbor model with k adjacent advantageous nodes, the transition from an absorbing state in $Z^{k+1}(M)$ to a state in $Z^{k+1}(M+2)$ can be induced via single mutations, if at least one of the clusters in z has no fewer than $(3k+2)$ players.

Proof: The proof of the lemma is analogous to Lemma 6. Consider an absorbing state z in $Z_k^{k+1}(M)$, in which (1) the k advantageous nodes are occupied by B-players, (2) at least$(k+1)$ B-neighbors on the left hand side of the advantageous nodes and (3) at least $(k+1)$ B-neighbors on the right hand side of the advantageous nodes,

$$z = \underbrace{BB\ldots BB}_{m_1'}\underbrace{B^*B^*\ldots B^*\mathbf{B}^*}_{k}\underbrace{BB\ldots BB}_{m_1''}\underbrace{\mathbf{A}A\ldots AA}_{m_2}\cdots\underbrace{BB\ldots BB}_{m_{M-1}}\underbrace{AA\ldots AA}_{m_M},$$

i.e. $m_1' + k + m_1'' = m_1 \geqslant 3k+2$, and $m_1', m_1'' \geqslant k+1$. In each step, one B-player located in the advantageous node swap with an A-player on the border of an A-cluster which has no less than $(k+2)$ players by mistake. After k steps, we will obtain a state z' via a series of single mutations:

$$z' = \underbrace{BB\ldots B\mathbf{B}}_{m_1'}\underbrace{A^*A^*\ldots A^*A^*}_{k}\underbrace{BB\ldots BB}_{m_1''}\underbrace{AA\ldots AA}_{m_2}\cdots\underbrace{BB\ldots BB}_{m_{M-1}}\underbrace{\mathbf{A}A\ldots AA}_{m_M}.$$

In the state z', we have $(M+2)$ clusters, every cluster containing normal nodes has no fewer than $(k+1)$ players, the A-cluster who occupied the advantageous nodes has k players. Apparently, z' belongs to $Z_k^{(k+1)'}$. In the next step, another A-player on the border of a cluster can swap to a B-player next to the advantageous nodes with one mutation and will be satisfied there. Now we reach an absorbing state z'':

$$z'' = \underbrace{BB\ldots BB}_{m'_1}\underbrace{AA^*A^*\ldots A^*A^*}_{k+1}\underbrace{BB\ldots BB}_{m''_1}\underbrace{AA\ldots AA}_{m_2}\cdots\underbrace{BB\ldots BB}_{m_{M-1}}\underbrace{AA\ldots AA}_{m_M}.$$

We find that there are $(M+2)$ clusters in z'' , each of the clusters has no fewer than $(k+1)$ players, so we have $z'' \in Z_k^{k+1}(M+2)$. Thus, when at least one of the clusters has $(3k+2)$ players, it is possible to move from an absorbing state in $Z_k^{k+1}(M)$ to an absorbing state in $Z_k^{k+1}(M+2)$ via a series of single mutations. $\square$

By Lemma 12, we know the condition under which we can increase the number of cluster in the absorbing states via single mutations. Next we characterize the LRE in $2k$-neighbor model with k adjacent advantageous nodes.

Proposition 3: Let r be the remainder when n is divided by $(k+1)$. In the $2k$-neighbor model with k adjacent advantageous nodes,

1) If $r = k$, the set of LRE includes all the absorbing states in Z_k^{k+1};

2) If $r < k$, the set of LRE includes all the absorbing states in $Z_k^{k+1}/Z_k^{k+1}(2\lfloor\frac{n}{k+1}\rfloor)$.

Proof: Consider an absorbing state z with $(2\lfloor\frac{n}{k+1}\rfloor-2)$ clusters, i.e. $z \in Z_k^{k+1}(2\lfloor\frac{n}{k+1}\rfloor-2)$. The largest possible cluster of z can have $(2k+2+r)$ players. By Lemma 12, we know that if z can reach states in $Z_k^{k+1}(2\lfloor\frac{n}{k+1}\rfloor)$ via single mutations, one of its clusters has to have no fewer than $(3k+2)$ players. In this case, the remainder r should be no less than k, i.e. $r \geqslant k$. By the definition of remainder, we know that $r < k+1$, and r is a natural number, so we have $r = k$.

Now we consider the following two cases.

a) When $r = k$, the largest possible cluster in a state in $Z_k^{k+1}(2\lfloor\frac{n}{k+1}\rfloor-2)$ has $(3k+2)$ players. By lemma 12, the states in $Z_k^{k+1}(2\lfloor\frac{n}{k+1}\rfloor)$ and the rest of the absorbing states in Z_k^{k+1} can reach each other via single mutations. By Lemma 11 and Lemma 12, all the absorbing states in Z_k^{k+1} are included in the same mutation-connected component.

b) When $r < k$, the largest possible cluster in a state in $Z_k^{k+1}(2\lfloor \frac{n}{k+1} \rfloor - 2)$ has fewer than $(3k+2)$ players. By lemma 12, the state in $Z_k^{k+1}(2\lfloor \frac{n}{k+1} \rfloor)$ cannot be reached from other states in Z_k^{k+1} via single mutations. By Lemma 11 and Lemma 12, the states in $Z_k^{k+1}/Z_k^{k+1}(2\lfloor \frac{n}{k+1} \rfloor)$ can reach each other via single mutations. Thus, the states in $Z_k^{k+1}/Z_k^{k+1}(2\lfloor \frac{n}{k+1} \rfloor)$ are included in the same mutation-connected component.

Next, we consider the transitions between the absorbing states in $Z_k^{(k+1)'}$ and the absorbing states in Z_k^{k+1}. We know that the only difference between the states in $Z_k^{(k+1)'}$ and the states in Z_k^{k+1} is the set of players in the advantageous nodes.

i) First, we show that, given every cluster containing the normal nodes has at least $(k+1)$ players, a state in Z_k^{k+1} can reach a state in $Z_k^{(k+1)'}$ via single mutations. To reach an absorbing state in $Z_k^{(k+1)'}$ from a state $z \in Z_k^{k+1}$, we need that there exists at least one advantageous node in z, such that (1) the players in the normal nodes will be satisfied regardless of which type of player is located in this advantageous node, (2) the advantageous node is not on the border of a cluster. If there exist such an advantageous node, consider the swap between the players located in the advantageous node and an opposite type player in the normal node, who is on the border of a cluster with at least $(k+2)$ players,

$$z = \underbrace{BB\dots BB}_{m_M}\underbrace{A^*\mathbf{A}^*\dots A^*A^*AA\dots AA}_{m_1}\underbrace{BB\dots B\mathbf{B}}_{m_2}\underbrace{AA\dots AA}_{m_3}\dots\underbrace{BB\dots BB}_{m_{M-2}}\underbrace{AA\dots AA}_{m_{M-1}}.$$

After the swap, we obtain an state z',

$$z' = \underbrace{BB\dots BB}_{m_M}\underbrace{A^*\mathbf{B}^*\dots A^*A^*AA\dots AA}_{m_1}\underbrace{BB\dots B}_{m_2-1}\underbrace{\mathbf{A}A\dots AA}_{m_3+1}\dots\underbrace{BB\dots BB}_{m_{M-2}}\underbrace{AA\dots AA}_{m_{M-1}}.$$

In z', every cluster containing the normal nodes has at least $(k+1)$ players, the B^*-player has fewer than k B-neighbors, so we know that $z' \in Z_k^{(k+1)'}$.

ii) Next, we show that a state in $Z_k^{(k+1)'}$ can reach a state in Z_k^{k+1} via single mutations. the transition from states in $Z_k^{(k+1)'}$ to states in Z_k^{k+1} is simply the transition in the opposite direction. In each step, choose a player in the advantageous node who has fewer than k neighbors of own type, and swap with a opposite type player on the border of a cluster. Iterating this process until every cluster in the city has at least $(k+1)$ players, then we will obtain a state in Z_k^{k+1}.

Thus, the transition between the states in $Z_k^{(k+1)'}$ and the states in Z_k^{k+1} can be induced via a series of single mutations.

Then, we consider the transition between the absorbing states in Z_k^m and the absorbing states in Z_k^{k+1}. As shown in Proposition 2, one mistake is enough to leave the basin of attraction of the states in Z_k^m. On the other hand, to induce the transition from an absorbing state in Z_k^{k+1} to an absorbing state in Z_k^m, we need to increase the number of clusters in the city. To increase the number of clusters via single mutations, we need at least one of the clusters has no fewer than $(3k+2)$ players, so that the states in $Z_k^{k+1}(2\lfloor \frac{n}{k+1} \rfloor)$ and the rest of the absorbing states in Z_k^{k+1} can reach each other via single mutations. However, states in Z_k^m have strictly more than $2\lfloor \frac{n}{k+1} \rfloor$ clusters. Thus, we cannot induce a transition from an absorbing state in Z_k^{k+1} to an absorbing state in Z_k^m via single mutations, which implies that the transition cost is strictly larger than one.

Finally, we consider the set of LRE. Let L denote the total number of absorbing states in the model. Recall that when $r = k$, all the absorbing states in Z_k^{k+1} form a mutation-connected component. All the z-trees directing to the states in Z_k^{k+1} have the least resistance of $L-1$, and there does not exist any other absorbing state, such that the z-tree directing to this state has the least resistance which is strictly less than $L-1$. Thus, in this case, all the absorbing states in Z_k^{k+1} are included in the set of LRE. By the same argument, when $r < k$, all the states in $Z_k^{k+1}/Z_k^{k+1}(2\lfloor \frac{n}{k+1} \rfloor)$ are included in the set of LRE. □

Note that Proposition 1 is a special case of Proposition 4, when $k = 1$.

6 Simulations

The main question in this section is whether the presence of an advantageous node can lead us to a more integrated society, when the parameter values are finite. In our simulation study, we replicate the theoretical model in which players of two types locate in a one-dimensional circular city with one advantageous node. As a benchmark, we also consider the case where the nodes in the city are homogeneous, and this is equivalent to the theoretical model in Young (1998). We refer to the model with one advantageous node as $M1$, and refer to the model of homogeneous nodes as $M2$.

To measure the degree of segregation, we count the number of clusters in the city, which is also employed in Schelling (1969, 1971a, 1971b, 1978). A state with more clusters represents a more integrated city, and in a completely segregated state there are only 2 clusters.

		$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.001$	$\epsilon = 0.0001$
$n = 10$	$M1$	3.18	2.52	2.48	2.52
	$M2$	3.02	2.25	2.15	2.24
$n = 20$	$M1$	5.62	3.69	3.33	3.89
	$M2$	5.48	3.33	2.93	3.45
$n = 40$	$M1$	10.52	6.16	5.41	7.14
	$M2$	10.42	5.77	4.58	6.24

Table 1: Average cluster number

We run 100 simulations with 100000 periods for each model. All the simulations start from a random state, in which players of two types are randomly distributed in the circular city. In each period, one pair of players is drawn at random. If the pair can make a Pareto-improving swap, they will exchange their positions with probability $(1-\epsilon)$, and stay where they are with probability ϵ. If the pair is not Pareto-improved after exchanging their locations, they will stay with probability $(1-\epsilon)$, and swap with probability ϵ.

Table 1 shows the average number of clusters for both models with different population sizes and error rates. We can find that $M1$ have more clusters than $M2$ in all cases. So the main conclusion of our theoretical model holds, the advantageous node do lead to a more integrated society. We can also have some insights into the results with comparative static analysis.

First, given error rate, the average number of clusters increases with the population size. The reason is as the population size increases, the pair of players which can make Pareto-improving swap will be drawn with lower probability. Consider a state z with 4 clusters, one of which has only one A-player.

$$z = \underbrace{A \ldots A \vdots A^* A \ldots A}_{m_1} \underbrace{\mathbf{B} \ldots B}_{m_i} \underbrace{\mathbf{A}}_{1} \underbrace{B \ldots \mathbf{B}}_{m_M} \underbrace{A \ldots A \vdots A^* A \ldots A}_{m_1}.$$

This unsatisfied player can make a Pareto-improving swap with two B-players (Both B-players are indicated in bold). When the population size is ten, the probability that the isolated A-players and one of the emboldened B-players are chosen is $2/C_2^9 \approx 0.05$; when the population size is 40, the probability reduce to $2/C_2^{39} \approx 0.002$.

Second, given the population size, the average number of clusters increases with the error rate. This observation is due to the conflict between the population size and error term. In the stochastical stability analysis, we assume that an error occurs with probability ϵ, which approaches 0. However, as we mentioned above, the undisturbed dynamics in our model will take place with lower probability as the number of players increases. To ensure that the error term will correctly reflect the long-run behavior of the dynamics, it should be extremely small.

	$\epsilon = 0.1$	$\epsilon = 0.01$	$\epsilon = 0.001$	$\epsilon = 0.0001$
$n = 10$	5.2%	11.6%	15%	12%
$n = 20$	2.6%	10.7%	13.8%	13%
$n = 40$	1%	6%	18.1%	14%

Table 2: Relative increase in cluster number

Therefore, the simulation gives a more significant result in the cases where the error rate is small enough. Table 2 summarizes the relative increase in cluster number from $M2$ to $M1$. We find that when $\epsilon = 0.001$, the effect of the advantageous node is most significant. This is because the mistakes in the model not only provide the opportunity of transitions between absorbing states, but also provide the opportunity of the transitions from absorbing state to transient state. To reach LRE from mid-term equilibria in finite periods, the error rate should not be too small; on the other hand, the error rate should be restricted to prevent the transition from absorbing state to transient state occurring frequently.

7 Conclusion

We have shown that the result of the standard Schelling model is not robust with respect to location heterogeneity. In particular, if there are locations such that an agent residing there is impartial to the composition of her neighborhood, integrated profile may also emerge as LRE. When the neighborhood size is larger than two, then more than one advantageous nodes with that property may be needed to obtain this result.

References

Bankston, C. & Caldas, S. J. (1996), 'Majority african american schools and social injustice: The influence of de facto segregation on academic achievement', *Social Forces* 75(2), 535–555.

Benard, S. & Willer, R. (2007), 'A wealth and status-based model of residential segregation', *Mathematical Sociology* 31(2), 149–174.

Bruch, E. E. & Mare, R. D. (2006), 'Neighborhood choice and neighborhood change1', *American Journal of Sociology* 112(3), 667–709.

Charles, C. Z., Dinwiddie, G. & Massey, D. S. (2004), 'The continuing consequences of segregation: Family stress and college academic performance*', *Social Science Quarterly* 85(5), 1353– 1373.

Clark, W. A. (1991), 'Residential preferences and neighborhood racial segregation: A test of the schelling segregation model', *Demography* 28(1), 1–19.

Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D. & York, R. (1966), 'Equality of educational opportunity', Washington, dc pp. 1066–5684.

Collins, C. A. &Williams, D. R. (1999), Segregation and mortality: the deadly effects of racism?, in 'Sociological Forum', Vol. 14, Springer, pp. 495–523.

Cutler, D. M. & Glaeser, E. L. (1997), 'Are ghettos good or bad?', *Quarterly Journal of Economics* 112(3).

Cutler, D. M., Glaeser, E. L. & Vigdor, J. L. (1997), The rise and decline of the american ghetto, Technical report, *National Bureau of Economic Research.*

Duncan, O. D. & Duncan, B. (1957), The Negro population of Chicago: A study of residential succession, *University of Chicago Press.*

Ellison, G. (2000), 'Basins of attraction, long-run stochastic stability, and the speed of step-bystep evolution', The *Review of Economic Studies* 67(1), 17–45.

Epstein, J. M. & Axtell, R. (1996), Growing artificial societies: social science from the bottom up, Brookings Institution Press.

Farley, R., Fielding, E. L. & Krysan, M. (1997), 'The residential preferences of blacks and whites: A four-metropolis analysis', *Housing Policy Debate* 8(4), 763–800.

Fossett, M. & Waren, W. (2005), 'Overlooked implications of ethnic preferences for residential segregation in agent-based models', *Urban Studies* 42(11), 1893–1917.

Freidlin, M. & Wentzell, A. (1988), Random Perturbations of Dynamical Systems, 2nd Ed., Springer Verlag, New York.

Jencks, C. & Mayer, S. E. (1990), 'Residential segregation, job proximity, and black job opportunities', Inner-city poverty in the United States pp. 187–222.

Krivo, L. J. & Peterson, R. D. (1996), 'Extremely disadvantaged neighborhoods and urban crime', *Social forces* 75(2), 619–648.

Laurie, A. J. & Jaggi, N. K. (2003), 'Role of'vision'in neighbourhood racial segregation: a variant of the schelling segregation model', *Urban Studies* 40(13), 2687–2704.

Massey, D. S. & Denton, N. A. (1993), American apartheid: Segregation and the making of the underclass, Harvard University Press.

Massey, D. S. & Fischer, M. J. (2006), 'The effect of childhood segregation on minority academic performance at selective colleges', *Ethnic and Racial Studies* 29(1), 1–26.

Mayer, S. E. (2002), 'How economic segregation affects children's educational attainment', *Social forces* 81(1), 153–176.

Nöldeke, G. & Samuelson, L. (1993), 'An Evolutionary Analysis of Backward and Forward Induction', *Games and Economic Behavior* 5, 425–454.

O'Sullivan, A. (2009), 'Schelling¡ä s model revisited: Residential sorting with competitive bidding for land', *Regional Science and Urban Economics* 39(4), 397–408.

Pancs, R. & Vriend, N. J. (2007), 'Schelling's spatial proximity model of segregation revisited', *Journal of Public Economics* 91(1), 1–24.

Schelling, T. C. (1969), 'Models of segregation', The *American Economic Review* pp. 488–493.

Schelling, T. C. (1971a), 'Dynamic models of segregation?', *Journal of mathematical sociology* 1(2), 143–186.

Schelling, T. C. (1971b), 'On the ecology of micromotives', The *Public Interest* (25), 59.

Schelling, T. C. (1972), 'A process of residential segregation: neighborhood tipping', Racial discrimination in economic life 157, 174.

Schelling, T. C. (1978), Micromotives and macrobehavior, WW Norton & Company.

Shihadeh, E. S. & Flynn, N. (1996), 'Segregation and crime: The effect of black social isolation on the rates of black urban violence', *Social Forces* 74(4), 1325–1352.

Taeuber, K. E. (1965), Negroes in cities.

Wilson, W. J. (2012), The truly disadvantaged: The inner city, the underclass, and public policy, University of Chicago Press.

Young, H. P. (1998), Individual strategy and social structure: An evolutionary theory of institutions, Princeton University Press.

Zhang, J. (2004a), 'A dynamic model of residential segregation', *Journal of Mathematical Sociology* 28(3), 147–170.

Zhang, J. (2004b), 'Residential segregation in an all-integrationist world', *Journal of Economic Behavior & Organization* 54(4), 533–550.

Social Coordination and Network Formation in Two Groups

Abstract

In this paper, we consider a model of social coordination and network formation, where players of two groups play a 2×2 coordination game when they are connected. Players in one group actively decide on whom they play with and on the action in the game, while players in the other group decide on the action in the game only and passively accept all the connections from other group. The players in the active group can connect to a limit number of opponents in the other group. We find that the selection of long-run outcome is determined by the population size of the groups, not the overall population size of them. If either group's population size is small in comparison to the linking restriction, all players will choose the risk-dominant equilibrium, while when both groups are sufficiently large in population, the players of two groups will coordinate on the payoff dominant action.

1 Introduction

In many social and economic interactions, agents can benefit from coordination on the same strategies or common standards. Many researchers have explored the interactions within the same population.[15] Nevertheless, a lot of interactions feature different groups of agents, with interactions only happening across groups. Think, for instance, of vertical relationship in a retail industry, where a retailer may more frequently interact with manufacturers located upstream, rather than with his horizontal competitors. It is reasonable to expect that rather than interacting with all manufacturers, the interaction will be described by a social network. Moreover, due to the market structures or the price elasticities in different industries, the relationships between the upstream firms and the downstream firms may be not symmetric. Consider the example of the retail industry again. It is more likely that the retail giants, e.g. Wal-Mart, Tesco, Carrefour, actively decide on their interaction partners among the manufacturers, whereas the manufacturers are in the passive position against these retail giants.

We aim to understand the implications of the action choices and interaction structures among the two groups in a setting where the interaction structure is the one-sided decision of one group. We present a simple model of action choice and network formation, encompassing two groups of players, called the M-group and the F-group. The players in the M-group decide on the action and the set of interaction partners, while the players in the F-group decide on the action and accept all the links from the M-players. An M-player and an F-player play a 2×2 coordination game when they are connected. In each period, players of two groups choose links and/or actions to maximize (myopically) their respective payoffs. We are interested in the scenario where the M-players can sustain a limited number of links à la Staudigl and Weidenholzer (2014). In this case, the M-players might not connect to all the F-players.

[15] For global interaction models see, e.g., Kandori et al. (1993), Kandori & Rob (1995), Young (1993). For local interaction models see Blume (1993, 1995), Ellison (1993, 2000) and Alós-Ferrer and Weidenholzer (2008); see also Weidenholzer (2010) for a recent survey on local interaction models focusing on social coordination.

We postulate that if an F-player has no incoming link from M-players, she behaves as if she plays against the entire population of M-players.

To understand the dynamics of the model, we first characterize its static behavior. We find that there are two Nash equilibria: the risk-dominant convention where all players in both groups choose the risk-dominant action, and the payoff dominant convention where all players in both groups choose the payoff dominant action. To give a long run prediction, we assume that players occasionally make mistakes. Different from the previous works on network formation and social coordination, we have two groups of players in the model. The equilibrium selection is determined by the population sizes of both groups, not the overall population of the two groups. Only when both populations are large in comparison to the number of links that may be sustained, the players of two groups will coordinate on the payoff dominant action. On the contrary, if either of the two groups is small, even though the overall population of the two groups is large, all players will choose the risk-dominant equilibrium.

The intuition underlying the main results is as follows. Due to the coordination of the game, if all players in one group choose the same action, the other group will converge to the same action with positive probability. The transition from one convention to another has two paths. In one path, every M-player switches to a new action first, and F-players converge to that action in the following periods. In the other path, every F-player switches to a new action first, and the M-players follow by coordination. The transition is determined by the path with the fewest mistakes.

We examine the following two situations. In the first situation, every M-player can link to all F-players. The transitions between the conventions are determined by the path where the smaller group switch first, and the risk-dominant action will spread to the two groups in the long run. This is because in this situation, there is only one interaction structure among the groups. Under such a fixed interaction structure, the risk-dominant action has a larger basin of attraction. In the second situation, the M-players can only link to a fraction of F-players due to the linking constraint.

Thus, the M-players can choose different interaction structures in the conventions. The interaction structures among the groups will influence the transitions among the conventions. The necessary number of mutations to reach the payoff dominant convention is independent of the population sizes of both groups. While the transition from the payoff dominant convention to the risk-dominant convention needs a fraction of players in the smaller group to make mistakes. The risk-dominant convention will be easier to reach and more difficult to leave when the population sizes of both groups are small. When the populations of both groups are large in comparison to the linking constraint, the payoff dominant action will be selected in the long run.

The remainder of this paper is structured as follows. The related previous literature will be reviewed in the next section. Section 3 presents the model and discusses the main techniques used. Section 4 presents the main result. Section 5 concludes.

2 Literature Review

There is a branch of literature considers scenarios where in addition to the action choice in a game, the agent can select the set of interaction partners. In Jackson and Watts (2002), interaction requires the consent of both parties. The linking cost has a significant impact on the selection of equilibrium: the risk-dominant convention is selected when the linking cost is low; the risk-dominant convention and the payoff dominant convention are both selected when the linking cost is high. In Goyal and Vega-Redondo (2005), a player connects with other players by unilaterally investing in costly pairwise links. They show that if the cost of maintaining a link is relatively low, players will coordinate on the risk-dominant action, and will coordinate on the efficient action otherwise. Hojman and Szeidl (2006) present a model where players pay for their out-degree links, and receive payoffs from all path-connected neighbors. Bilancini and Boncinelli (2014) present a model of social coordination in a population made of two different types, players have a preference for own-type but can observe others' type only after first interaction. They find that the selection of conventions in the long run depends on the cost of mismatch in type. When the cost of mismatch in type is small with respect to the cost of mismatch in action, the unique long run outcome is the payoff dominant convention; when the cost of mismatch in type is large, both conventions coexist in the long run. Staudigl and Weidenholzer (2014) consider a model of social coordination and network formation with constrained interaction. They show that if this constraint of links is relatively small with respect to the population size, the payoff dominant convention will emerge in the long run. The main difference between our model and the previous works is that in our model, interactions take place across the groups. The M-players support links to the F-players, who accept all the incoming links. Agents on both sides of the link receive the payoff, only the M-players pay the cost of supporting the links.

A different branch of literature studies the models where in addition to the action choice in a game, the agents choose their interaction partners by choosing among several locations. This branch of literature includes Ely (2002), Oechssler (1997), Bhaskar and Vega-Redondo (2004). In these models, players can choose the location to play the game, which makes the player can easily avoid miscoordination by "voting by feet". Thus, players will choose the payoff dominant action in the long run. Anwar (2002) present multiple location models where each of the locations has a capacity constraint and some agents are immobile. The constraint will limit the movement between locations, in this context the payoff dominant convention will not be selected.

The present paper is also related to the literature in the study of matching markets. In the early literature (Gale and Shapley 1962, Roth and Sotomayor 1989), players match in a two-sided market, like labor markets or marriage market, but there's no further interaction after the matching. Later Kranton and Minehart (2003) introduce a buyer-seller network, and find that when buyers and sellers can form links strategically and compete in the matching markets, the efficient network structure is always an equilibrium outcome. Jackson and Watts (2008) generalize the matching problems by presenting the bipartite games, where players choose both a strategy in a game and a partner with whom to play the game. They prove the existence of equilibria in the bipartite games. In a companion paper, Jackson and Watts (2010) consider games played among multiple groups of players. As in Jackson and Watts (2008), players in the game choose their strategies as well as their interaction partners across the groups, those who are dissatisfied will rematch and reach a new equilibrium. They show that the long-run outcome depends on the relative sizes of population in different groups. In Pongou and Serrano (2013), men and women form a fidelity network by having relationships with the opposite type. They show that different cultures of gender relationship lead to different long-run networks. When women are easier to change their partners, all agents have the same number of partners in the long run. Otherwise, all women will match with a small fraction of men. Different from these works in this branch of literature, in our paper, we focus on the coordination game between the two groups. Moreover, in the matching process of our paper, the relationship between the two groups is asymmetric, players in the M-group makes the decision of interaction structure unilaterally, and F-players accept all linking invitation from M-players.

3 The Model

We consider two groups of players, called the M-players and the F-players. We denote by $M = \{1, 2, \ldots m\}$ the set of M-players, and denote by $F = \{m+1, m+2, \ldots m+f\}$ the set of F-players. Each $i \in M$ chooses a subset of F-players with whom to play a fixed bilateral game. Formally, we have $g_{ij} \in \{1, 0\}$ for any two players $i \in M$, $j \in F$. We say that $g_{ij} = 1$ when player i forms a link to player j, otherwise $g_{ij} = 0$. Let $g_i = (g_{i(m+1)}, g_{i(m+2)}, \ldots g_{i(m+f)})$ be the link formation choice of an M-player i. There is no direct link between players in the same set, i.e. for any two players $i, j \in M$, $g_{ij} = g_{ji} = 0$. A profile of link formation choices, one for each M-player is denoted by $g = (g_1, g_2 \ldots, g_m)$. We refer to g as the network of interaction. For each player $i \in M \cup F$, let $N(i)$ be the set of players who are directly connected to i, we say $N(i)$ is player i's neighborhood. For an F-player i, we denote by d_i the in-degree of player i, i.e. the number of players she is passively linked to.

Two players from different groups play a coordination game when they are directly connected. Each player i can choose an action $a_i \in \{A, B\}$. We denote by $u(a_i, a_j)$ the payoff of player i choosing action a_i given that player j chooses action a_j. The payoffs in the coordination game are given in the following matrix:

	A	B
A	a, a	c, d
B	d, c	b, b

We assume that $a > d$ and $b > c$, so that (A, A) and (B, B) are Nash equilibria. We assume that $b > a$, so (B, B) is the payoff dominant equilibrium. We assume that $a + c > b + d$, so the equilibrium (A, A) is risk-dominant in the sense of Harsanyi and Selten (1988), i.e. A is the unique best response against an opponent playing both strategies with equal probability. We denote by $q^* = \frac{b-c}{a+b-c-d}$, the Nash equilibrium (A, A) is risk-dominant implies that $q^* < \frac{1}{2}$. We assume that payoff flows to both sides of a link. We also assume that the payoffs in the bilateral game are all positive, i.e. $a, b, c, d > 0$, so that F-players have no incentive to refuse links from M-players. Let $a^M = (a_1^M, a_2^M \ldots a_m^M)$ be the profile of action choice of M-players and let $a^F = (a_1^F, a_2^F, \ldots a_f^F)$ be the profile of action choice of F-players.

For every M-player, in addition to her action choice in the coordination game, she also decides on which F-players to link to. We focus on a scenario in which every M-player can support at most ℓ links, where $\ell \leqslant f$ and $\ell \in \mathbb{N}$, where $\mathbb{N}$ denotes the natural number. There is at most one link between every pair of players. We denote the cost of sustaining a link by k. On the other hand, F-players offer no links but accept all the linking invitations from M-players. There is no restriction in the number of incoming links that F-players may receive.

A pure strategy of an M-player consists of her choice of action and the set of players she wants to link to. The set of linking strategies of player i is denoted by $\mathcal{G}_i = \{0,1\}^f$. The set $\mathcal{G} = \mathcal{G}_1 \times \mathcal{G}_2 \times \cdots \times \mathcal{G}_m$ is the space of pure linking strategies of all M-players. We denote the strategy profile of an M-player i by $s_i^M = (a_i, g_i) \in S_i^M = \{A, B\} \times \mathcal{G}_i$, let $s^M = (s_1^M, \dots s_m^M) \in S^M = \prod_{i \in M} S_i^M$ be the profile of strategies of the M-group; a pure strategy of an F-player j is simply her choice of action in the coordination game $s_j^F = a_j \in S_j^F = \{A, B\}$. The profile of strategies of the F-group is $s^F = a^F \in S^F = \prod_{i \in F} S_i^F$.

The overall payoff of a player is determined by the sum of payoffs she receives from the coordination games net of the linking cost. Given the strategy profile of M-player i, $s_i^M = (a_i, g_i)$, and the strategies of other players, $s_{-i} = (s_1^M, \dots s_{i-1}^M, s_{i+1}^M, \dots, s_m^M, s^F)$, her overall payoff is given by:

$$U_i^M(s_i^M, s_{-i}) = \sum_{j \in F} g_{ij} u(a_i, a_j) - k \sum_{j \in F} g_{ij}$$

Given the strategies of other players, $s_{-i} = (s^M, s_1^F \dots s_{i-1}^F, s_{i+1}^F, \dots s_{m+f}^F)$, the overall payoff of a player $i \in F$ from playing $s_i^F = a_i$ is:

$$U_i^F(s_i^F, s_{-i}) = \sum_{j \in M} g_{ji} u(a_i, a_j) \ + \frac{p}{M} \sum_{j \in M} u(a_i, a_j),$$

where $p = 0$ if $\sum_{j \in M} g_{ji} > 0$, and $p \in (0, \frac{d}{b})$ if $\sum_{j \in M} g_{ji} = 0$. The first term of the payoff function is the payoffs from coordination games with her neighbors when she has any incoming links from M-players, the second term is her payoff from interacting with the whole M-group, here we assume that p is sufficiently small so that we have $u(a_i, a_j) > p \sum_{j \in M} u(a_i, a_j)$ for any a_i, a_j. According to this payoff function, an unlinked F-player receives a lower payoff from non-connected interaction with the whole M-group, while if an M-player is connected to some M-players, the payoff from coordination games with her neighbors is always higher than that from non-connected interaction. Note that this assumption implies that when an F-player does not have any incoming links, she behaves as if she plays against the entire population of M-players, i.e. we postulate that she plays the field when she is unlinked. On one hand, this assumption may reflect the expectation that in the future the F-player will be connected to some M-player; on the other hand, from a more technical perspective, it avoids arbitrary behavior of F-players, which may determine the prediction of the model. See the discussion in footnote.[16]

In the following, we denote by $s = (s^M, s^F) = (a^M, g, a^F) \in S = S^M \times S^F$ the state in which the action profile of M-players is a^M, the action profile of F-players is a^F, and g represents the linking decision of M-players. We denote by m^A and f^A the number of players choosing action A in the M-group and the F-group respectively. It follows that the number of B-players in the F-group is $(f - f^A)$, and the number of B-players in the M-group is $(m - m^A)$. We further denote by $\overrightarrow{A}$ the action vector in which every player in a certain group adopts the same action A, and by $\overrightarrow{B}$ the action vector in which every player in a certain group adopts the same action B, e.g. we have $a^M = \overrightarrow{A}$ if every M-player chooses action A.

Time is discrete, denoted by $t = 1, 2, 3, \cdots$. At each period t, a state is given by the strategy profile of M-players and F-players, specifying the action played and the links established. At each period t, every player may revise her strategy with exogenous probability $\lambda \in (0, 1)$. When such a revision opportunity arises, she chooses a myopic best response to the other players' strategies in the preceding period. Formally, in period t, player i chooses the strategy:

[16] In this present model, the actions of M-players are determined by the distribution of the F-group. Due to random tie breaking in case of payoff ties, if the unlinked F-players can arbitrarily change their action, the optimal action for M-players may change without mutation. This will turn out that the payoff dominant equilibrium will be LRE in most cases.

$$s_i(t) \in arg\max_{s_i \in S_i} U_i(s_i, s_{-i}(t-1)),$$

where $s_{-i}(t-1)$ represents the strategy profiles of the other players except i in the previous period. If a player has multiple best replies, we assume that she randomly chooses one of them with exogenously given probability. We assume that with probability $\epsilon \in (0,1)$, an updating player makes mistakes and simply picks a strategy – consisting of action and/or linking choice – at random. We assume that ϵ is independent across players, time and payoffs. We refer to the process without mistakes ($\epsilon = 0$) as the unperturbed process and refer to the process with mistakes ($\epsilon > 0$) as the perturbed process.

3.1 Review of Techniques

We have denoted by s a state of the model, which specifies how M-players and F-players are connected and choose their actions in graph g, and we have denoted by S the set of states. An absorbing state is a state in which no alternative state s' can be reached from s without mutations.

The perturbed process is ergodic, i.e. it has a unique invariant distribution $\mu(\epsilon)$, which summarizes the long-run behavior of the process, independently of initial conditions. The limit invariant distribution (as the rate of experimentations tends to zero) $\mu^* = \lim_{\epsilon \to 0} \mu(\epsilon)$ exists and is an invariant distribution of the process without mistakes. The limit invariant distribution singles out a stable prediction of the process without mistakes ($\epsilon = 0$), in the sense that, for any ϵ small enough, the play approximates that described by μ^* in the long run. The set of states S that supports μ^* are called stochastically stable states or long-run equilibria (LRE).

We will rely on the characterization of the set of LRE developed by Ellison (2000). Given two absorbing set X and Y, denote $c(X, Y)$ as the minimal number of mutations necessary for a direct transition from X to Y, this direct transition from X to

Y does not go through any other absorbing set, and $c(X,Y) > 0$. Define a path from X to Y as a finite sequence of absorbing sets $P = \{X = S_0, S_1, \ldots S_{L(P)} = Y\}$, where $L(P)$ is the length from X to Y, i.e. the number of elements of the sequence minus 1. Let $W(X,Y)$ be the set of all paths from X to Y. We extend the cost function to paths by $c(P) = \sum_{k=1}^{L(p)} c(S_{k-1}, S_k)$, then the minimal number of mistakes require for a transition, direct or indirect, from X to Y is given by:

$$C(X,Y) = \min_{P \in W(X,Y)} c(P).$$

The result of Ellison (2000) can be summarized as follow: The radius of an absorbing set X is defined as

$$R(X) = \min\{C(X,Y) \mid Y \text{ is an absorbing set, } Y \neq X\},$$

i.e. the minimal number of mistakes necessary for leaving X.

We define the coradius of X as the maximal number of mistakes necessary for every other absorbing set to enter the basin of attraction of X, formally:

$$CR(X) = \max\{C(Y,X) \mid Y \text{ is an absorbing set, } Y \neq X\}.$$

Ellison (2000) provides a powerful result that if $R(X) > CR(X)$ for a given absorbing set X, then X is the unique stochastically stable set.

Lemma 1 (Ellison, 2000) Let X be an absorbing set, if $R(X) > CR(X)$, then the LRE are the states in X.

This is the Radius-Coradius Theorem of Ellison (2000). Note that if there are only two absorbing set, we have $C(X,Y) = R(X) = CR(Y)$ and $C(Y,X) = R(Y) = CR(X)$; when $C(X,Y) = C(Y,X)$, both absorbing sets are LRE.

4 Coordination in Two Groups

Note that as the M-players choose their actions and links simultaneously, we can characterize the optimal behavior of an M-player by splitting his decision into two parts: First, to chooses the optimal set of links for both actions A and B, and second, to decide which action to adopt, given the optimal set of links.

Let us start by considering an M-player's optimal linking strategy. As we consider the case of low linking cost ($k < d$), every link is valuable for M-player regardless of her own action choice and her potential opponent's action. Thus, an M-player will always form all her ℓ links to F-players.

Note that, an A-player will prefer interacting with another A-player over interacting with a B-player. If an M-player chooses action A, her optimal linking strategy will be: First, to establish links to A-players in the F-group; second, if there are spare links after connecting to all possible A-players, to use the rest of her links connecting to B-players in the F-group. Similarly, an M-player choosing action B will optimally link up first to B-players in the F-group, and then link up to A-players if she has spare links.

Given the optimal linking strategy, the action choices of M-players depend on the distribution of action in the F-group (a^F). We denote by $V(T, a^F)$ the maximal payoff that an M-player with action $T \in \{A, B\}$ can obtain from linking up optimally, given a^F. So we have

$$V(A, f^A) = a \min\{\ell, f^A\} + c(\ell - \min\{\ell, f^A\})$$

and

$$V(B, f^A) = d(\ell - \min\{\ell, f - f^A\}) + b \min\{\ell, f - f^A\}$$

An M-player will choose A with probability one if $V(A, f^A) > V(B, f^A)$, will choose B with probability one if $V(A, f^A) < V(B, f^A)$, and will randomize when $V(A, f^A) =$

$V(B, f^A)$.

Depending on the relationship between ℓ, f and f^A, we have four cases:

(i) $f^A \geqslant \ell$ and $f - f^A \geqslant \ell$

In this case, both A-players and B-players may fill up all their links with F-players of their own kind. Since (B, B) is the payoff dominant equilibrium, M-players will earn a higher payoff by adopting B, so B is the optimal action in this case.

(ii) $f^A < \ell$ and $f - f^A \geqslant \ell$

In this case, B-players can fill up all their links to F-players of their own kind, but A-players do not find sufficiently many F-players of their own kind to fill up all the slots. Again B is the optimal action in this case.

(iii) $f^A \geqslant \ell$ and $f - f^A < \ell$

In this case, A-players can fill up all their links to F-players of their own kind, but B-players cannot. An M-player will choose A if $V(A, f^A) > V(B, f^A)$, which turns to

$$a\ell > d(\ell - f + f^A) + b(f - f^A),$$

rearranging terms yields

$$f^A > f - \frac{(a-d)\ell}{b-d}.$$

Conversely, note that if when $f^A < f - \frac{(a-d)\ell}{b-d}$, $V(B, f^A) > \max V(A, f^A) = a\ell$, i.e. if the number of B-players in F-group is larger than $\frac{(a-d)\ell}{b-d}$, it is always optimal for M-players to play B. We denote by $\psi_1 = \frac{(a-d)\ell}{b-d}$.

(iv) $f^A < \ell$ and $f - f^A < \ell$

In this case, neither A- nor B-players in the M-group will fill up all their links with F-players of their own kind. The M-players will choose A with probability one if

$$f^A > \frac{(b-d)f-(c-d)\ell}{a+b-c-d}.$$

We denote by $\psi_2 = f - \frac{(b-d)f-(c-d)\ell}{a+b-c-d}$.

We summarize the conditions under which players choose either of the two actions in the following table:

Table 1 **Action Choice Thresholds**

	$f^A \geqslant \ell$ $f-f^A \geqslant \ell$	$f^A < \ell$ $f-f^A \geqslant \ell$	$f^A \geqslant \ell$ $f-f^A < \ell$	$f^A < \ell$ $f-f^A < \ell$
Choose A with prob. 1	Never	Never	$f^A > f-\psi_1$	$f^A > f-\psi_2$
Choose B with prob. 1	Always	Always	$f^A < f-\psi_1$	$f^A < f-\psi_2$
Randomize	Never	Never	$f^A = f-\psi_1$	$f^A = f-\psi_2$

Next we consider the optimal strategy of F-players. For an F-player i, let d_i^A be the number of A-players in player i's neighborhood, let d_i^B be the number of B-players in player i's neighborhood. Formally, we have $d_i^A = \#\{j \in N(i) \mid a_j = A\}$ and $d_i^B = \#\{j \in N(i) \mid a_j = B\}$.[17] Recall that there are two kinds of F-players: Those who are connected to at least one M-player ($\sum_{j\in M} g_{ji} > 0$), and those who are not ($\sum_{j\in M} g_{ji} = 0$).

(1) When $\sum_{j\in M} g_{ji} > 0$, the action of an F-player i depends on d_i^A. When $d_i^A > q^* d_i$, she will choose action A with probability one, when $d_i^A < q^* d_i$, she will choose action B with probability one, and when $d_i^A = q^* d_i$, she will randomize between A and B.

(2) When $\sum_{j\in M} g_{ji} = 0$, the action of an F-player i depends on a^M. When $m^A > q^* m$, she will choose action A with probability one, when $m^A < q^* m$, she will choose action B with probability one, and when $m^A = q^* m$, she will randomize between A and B.

[17] We define $\#\{X\}$ to be the cardinality of a set X.

Before we characterize the set of absorbing states, we denote by $\overrightarrow{A}_g = \{s \in S \mid a_i = A,\ \forall i \in M \cup F \text{ and } d_i = \ell,\ \forall i \in M\}$ the set of states where all players choose action A, and the M-players link up with the maximal number of F-players possible, $\overrightarrow{B}_g$ is defined accordingly. Note that if $f > \ell$, M-players have more potential interaction partners than available links. Since the M-players will randomize among their strategies in case they are indifferent, the process will move among different interaction structures without mistake. This implies that the absorbing sets contain more than one element.

Lemma 2: The sets $\overrightarrow{A}_g$ and $\overrightarrow{B}_g$ are the only absorbing sets.

Proof: First note that since every connection carries a positive net-payoff, M-players will use up all their links. From the analysis above, we find that if all players in one group choose the same action, it is optimal for all players in the other group to choose that action. Consider a state $s \in \overrightarrow{A}_g \cup \overrightarrow{B}_g$, a revising player will always remain at her current action in the following period. Furthermore, since we assume that ties are broken randomly in case of payoff ties, for each pair of states $s_i, s_j \in \overrightarrow{A}_g$ (and also for each pair in $\overrightarrow{B}_g$), there is a positive probability of moving between them without mutation. It follows that states in $\overrightarrow{A}_g$ forms an absorbing set, and states in $\overrightarrow{B}_g$ forms an absorbing set.

Next, consider any state $s \notin \overrightarrow{A}_g \cup \overrightarrow{B}_g$. Consider the M-players, with positive probability, they will all choose the same action, (since they all facing the same action distribution among F-players (a^F)). Thus, we can move to a state where $a^M = \overrightarrow{A}$ or $a^M = \overrightarrow{B}$ with positive probability. After that, since all the F-players face the M-group, in which all of them choose the same action, it follows that they too all change that same action. $\square$

Before we proceed to characterize the set of Long-run Equilibria, we discuss the transitions between the absorbing sets. As shown in Lemma 2, if all players in one group choose the same action, it is optimal for all players in the other group to choose that action. A state in which $a^M = \overrightarrow{T}$ or $a^F = \overrightarrow{T}$, where $T \in \{A, B\}$, lies in the basin of attraction of the absorbing set $\overrightarrow{T}_g$. Thus, for the transitions to $\overrightarrow{T}_g$, we have two paths: the path via states where $a^M = \overrightarrow{T}$ and the path via states where $a^F = \overrightarrow{T}$. Note that in the transition to $\overrightarrow{T}_g$ via states where $a^F = \overrightarrow{T}$, we need that every F-player will choose T with positive probability; in the transition to $\overrightarrow{T}_g$ via states where $a^M = \overrightarrow{T}$, we need that every M-player will choose T with positive probability. In both paths, the mutations may occur among the F-players and the M-players. We analyze two scenarios: one where the number of F-players equals the number of links that M-players can support ($f = \ell$), and one where the number of F-players is larger than the number of links that M-players can support ($f > \ell$).

When $f = \ell$ (see Figure 1 for instance), every M-player links to all F-players, which implies that every F-player has m incoming links. First, consider the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$. We want to calculate how many players need to switch in order to move from a state in $\overrightarrow{A}_g$ to a state in which either $a^F = \overrightarrow{B}$ or $a^M = \overrightarrow{B}$, so that the process will move to $\overrightarrow{B}_g$ with positive probability. We denote by $P_F(A, B)$ the minimal number of players who have to switch from action A to action B such that the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ occurs with positive probability via a state in which $a^F = \overrightarrow{B}$, and $P_M(A, B)$ is defined accordingly.

Note that if an F-player has no fewer than $m(1-q^*)$ M-players choosing B in her neighborhood, she will switch to B with positive probability. Thus, with $\lceil m(1-q^*)\rceil$ mistakes,[18] we can move to a state in which $a^F = \overrightarrow{B}$, from which we reach a state in $\overrightarrow{B}_g$ with positive probability. So we have $P_F(A, B) = \lceil m(1-q^*)\rceil$. Likewise, if an M-player has no fewer than $f(1-q^*)$ F-players choosing B in her neighborhood, she will switch to B with positive probability. With $\lceil f(1-q^*)\rceil$ mistakes, we can move to a state in which $a^M = \overrightarrow{B}$, and consequently reach a state in $\overrightarrow{B}_g$ with positive probability. Thus, we have $P_M(A, B) = \lceil f(1-q^*)\rceil$.

Next we denote by $P_M(B, A)$ the minimal number of players who switch from

[18]We define $\lceil x\rceil$ to be the smallest integer not less than x.

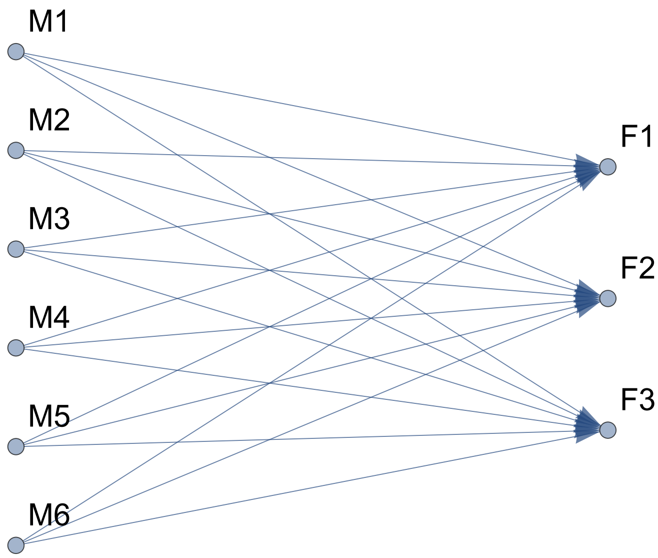

Figure 3: Interaction structure when $m = 6,\ f = \ell = 3$

action B to action A by mistake such that the transition from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$ occurs with positive probability via a state in which $a^M = \overrightarrow{B}$, and $P_F(B, A)$ is defined accordingly. With a similar analysis to the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$, it is straightforward to obtain that $P_F(B, A) = \lceil mq^* \rceil$.

Thus, when $f = \ell$, the transition cost between the absorbing sets are given by:

$$CR(\overrightarrow{B}_g) = R(\overrightarrow{A}_g) = \min\{P_M(A, B), P_F(A, B)\} = \lceil \min\{m, f\}(1 - q^*) \rceil$$

and

$$CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = \min\{P_M(B, A), P_F(B, A)\} = \lceil \min\{m, f\} q^* \rceil$$

When $f > \ell$, M-players have more potential interaction partners than available links. Since players are assumed to randomize among their strategies in case they are indifferent, the process will move among different interaction structures without mistake. (See Figure 2 for different interaction structures in an absorbing set.) Further, note that the transitions among different absorbing sets will depend on the particular interaction structures the process might visit. Recall that there are two paths from

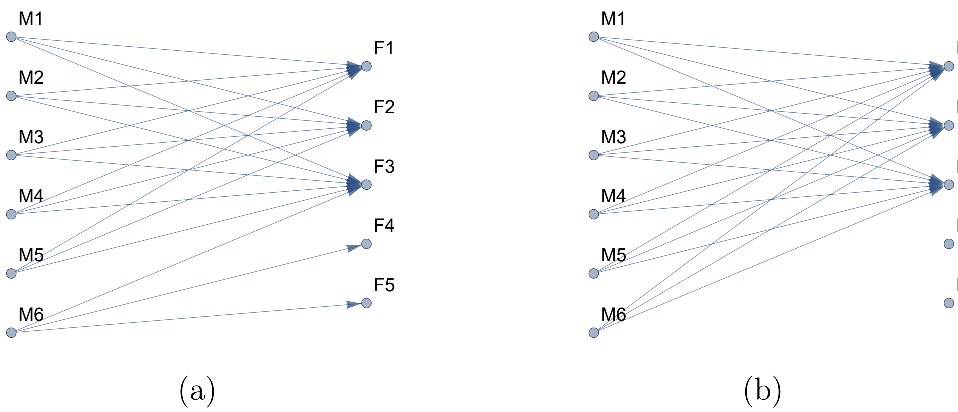

Figure 4: (a) An M-player influence structure when $m = 6$, $f = 5$, $\ell = 3$; (b) An F-player influence structure when $m = 6$, $f = 5$, $\ell = 3$

one absorbing set to another, one path via states where every M-player chooses the same action, and the other path via states where every F-player chooses the same action. In the following lemmas, we identify the transition cost of the two paths from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ and the two paths from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$.

Lemma 3 If $f > \ell$, the transition costs via states where every M-player chooses the same action are given as follow:

$$P_M(A, B) = \begin{cases} 1 & \text{if } f \geqslant \ell + \frac{(a-d)\ell}{b-d} \\ \lceil \psi_2 \rceil - (f - \ell) + 1 & \text{if } \ell < f < \ell + \frac{(a-d)\ell}{b-d} \end{cases}$$

and

$$P_M(B, A) = \begin{cases} \lceil f - \psi_1 \rceil - \ell + 1 & \text{if } f \geqslant 2\ell \\ \lceil f - \psi_1 \rceil - (f - \ell) + 1 & \text{if } 2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d} \\ \lceil f - \psi_2 \rceil - (f - \ell) + 1 & \text{if } \ell < f < \ell + \frac{(a-d)\ell}{b-d} \end{cases}.$$

The proof is given in the Appendix. Let us provide some technical intuition for the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states where $a^M = \overrightarrow{B}$. Similar arguments can be applied to the transition from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$ via states where $a^M = \overrightarrow{A}$. In the lemma, we find that interaction structures where one M-player can influence the highest number of F-players play an important role. We refer to such interaction structures as *M-player influence structures* (See Figure 2(a) for an example). Under such an interaction structure, $\min\{\ell, f-\ell\}$ F-players receive their only incoming link from one M-player. If this M-player changes, then also all of these F-players will change. If $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, following one M-player's mutation, more than $\lceil \psi_1 \rceil$ F-players switch to B, which in turn makes all the remaining M-players switch.[19] If $f < \ell + \frac{(a-d)\ell}{b-d}$, the M-player induces $(f-\ell)$ F-players to switch. In order for the remaining M-players to switch, we need other $(\lceil \psi_2 \rceil - (f-\ell))$ mutations among the F-players.

The following lemma characterizes the transition costs among absorbing sets via states where all F-players choose the same action.

Lemma 4 If $f > \ell$, the transition costs via states where every F-player chooses the same action are: $P_F(A, B) = \lceil m(1-q^*) \rceil$ and $P_F(B, A) = \lceil mq^* \rceil$.

The proof is given in the Appendix. Let us provide some technical intuition for the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states where $a^F = \overrightarrow{B}$. To understand the mechanism behind the transitions, consider an interaction structure where all the links of M-players are concentrated on ℓ F-players, and $(f-\ell)$ F-players are not connected. We refer to such interaction structures as *F-player influence structures* (See Figure 2(b) for an example). Under an F-player influence structure, when $\lceil m(1-q^*) \rceil$ M-players switch to B, both the F-players with incoming links and the F-players without incoming link will switch. In the lemma, we show that there does not exist an alternative interaction structure, under which all F-players change with strictly fewer mutations among M-players. Similar arguments can be applied to the transition from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$ via states where $a^F = \overrightarrow{A}$.

Combining Lemma 3 and Lemma 4, if $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, the transition costs between

[19] One M-player can influence $\min\{\ell, f-\ell\}$ F-players in an M-player influence structure. It is always true that $\min\{\ell, f-\ell\} > \lceil \psi_1 \rceil$ when $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, so in this case the transition cost from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states where $a^M = \overrightarrow{B}$ is one.

the absorbing sets are given by:

$$CR(\overrightarrow{B}_g) = R(\overrightarrow{A}_g) = \min\{P_M(A,B), P_F(A,B)\} = \min\{1, \lceil m(1-q^*) \rceil\}$$

and

$$CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = \min\{P_M(B,A), P_F(B,A)\} = \min\{\lceil f-\psi_1 \rceil - \min\{f-\ell, \ell\}+1, \lceil mq^* \rceil\}.$$

If $\ell < f < \ell + \frac{(a-d)\ell}{b-d}$, the transition costs between the absorbing sets are given by:

$$CR(\overrightarrow{B}_g) = R(\overrightarrow{A}_g) = \min\{P_M(A,B), P_F(A,B)\} = \min\{\lceil \psi_2 \rceil - (f-\ell)+1, \lceil m(1-q^*) \rceil\}$$

and

$$CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = \min\{P_M(B,A), P_F(B,A)\} = \min\{\lceil f-\psi_2 \rceil - (f-\ell)+1, \lceil mq^* \rceil\}.$$

4.1 Long-run Equilibria

Now we can identify the set of LRE ($\mathcal{S}$) using the Radius-Coradius Theorem (Lemma 1). Based on the analysis above, we will distinguish three cases: when $f = \ell$, $\ell < f < \ell + \frac{(a-d)\ell}{b-d}$ and $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$.

Proposition 1 When $f = \ell$, the sets of LRE are
$\mathcal{S} = \overrightarrow{A}_g$ if $\lceil \min\{m, f\} q^* \rceil < \lceil \min\{m, f\}(1-q^*) \rceil$;
$\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ if $\lceil \min\{m, f\} q^* \rceil = \lceil \min\{m, f\}(1-q^*) \rceil$.

The proof is given in the Appendix. In Proposition 1, we find that when the players of two groups are fully connected, the transition cost between $\overrightarrow{A}_g$ and $\overrightarrow{B}_g$ is determined by the population size of the smaller group, and the risk-dominant equilibrium is always selected. For, the transition from one absorbing set to another requires one of the two populations to change. If $f = \ell$, the number of mutations required will be a fraction of one population. Thus, if one is smaller than the other, the transition with the fewest mutations will involve the smaller population switching first. As by risk dominance, $q^* < \frac{1}{2}$, the risk-dominant equilibrium is always selected. Moreover, the payoff dominant equilibrium may be selected in addition to a risk-dominant equilibrium if q^* is close to $\frac{1}{2}$, and/or if the size of the smaller population is small.

In the following two propositions, we will identify the set of LRE when $f > \ell$.

Proposition 2 When $\ell < f < \ell + \frac{(a-d)\ell}{b-d}$, there exist two thresholds, $\underline{m}$ and $\overline{m}$, with $\underline{m} < \overline{m}$, such that

(1) If $m < \underline{m}$, the set of LRE is given by

$\mathcal{S} = \overrightarrow{A}_g$ if $\lceil mq^* \rceil < \lceil m(1-q^*) \rceil$ and $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ if $\lceil mq^* \rceil = \lceil m(1-q^*) \rceil$.

(2) If $m > \overline{m}$:

The set of LRE is $\mathcal{S} = \overrightarrow{A}_g$ under the following conditions: i) If $\psi_2 \in \mathbb{Z}$[20] and $\ell < f < \frac{2(c-d)\ell}{b+c-a-d}$; or ii) If $\psi_2 \notin \mathbb{Z}$, f is even, and $\ell < f < \frac{2(c-d)\ell}{b+c-a-d}$; or iii) If $\psi_2 \notin \mathbb{Z}$, f is odd, and $\ell < f < \frac{2(c-d)\ell-(a+b-c-d)}{b+c-a-d}$;

The set of LRE is $\mathcal{S} = \overrightarrow{B}_g$ under the following conditions: i) If $\psi_2 \in \mathbb{Z}$ and $\frac{2(c-d)\ell}{b+c-a-d} < f < \ell + \frac{(a-d)\ell}{b-d}$; or ii) If $\psi_2 \notin \mathbb{Z}$, f is even, and $\frac{2(c-d)\ell}{b+c-a-d} < f < \ell + \frac{(a-d)\ell}{b-d}$; or iii) If $\psi_2 \notin \mathbb{Z}$, f is odd, and $\frac{2(c-d)\ell+(a+b-c-d)}{b+c-a-d} \leqslant f < \ell + \frac{(a-d)\ell}{b-d}$;

The set of LRE is $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ under the following conditions: i) If $\psi_2 \in \mathbb{Z}$ and $f = \frac{2(c-d)\ell}{b+c-a-d} \in \mathbb{Z}$; or ii) If $\psi_2 \notin \mathbb{Z}$, f is odd, and $\frac{2(c-d)\ell-(a+b-c-d)}{b+c-a-d} \leqslant f < \frac{2(c-d)\ell+(a+b-c-d)}{b+c-a-d}$.

Proposition 3 When $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, the set of LRE is given by $\mathcal{S} = \overrightarrow{B}_g$ if $m > \frac{1}{q^*}$, $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ if $1 \leqslant m \leqslant \frac{1}{q^*}$.

The proof is given in the Appendix. Let us provide some technical intuition for the result of these two propositions. As we stated above, the transition from one

[20] We define $\mathbb{Z}$ to be the set of integers.

absorbing set to another requires one of the two populations to change to the other action, so there are two paths to induce the transition among the absorbing sets. The transition cost among the absorbing states are determined by the path with least mutations. When the population size of the M-group is sufficiently small, the path via states where every F-player chooses the same action needs fewer mutations. By Lemma 4, we know that the transitions via this path require the fewest mutations under an F-player influence structure, and the risk-dominant convention ($\overrightarrow{A}_g$) will always be selected. When the population size of the M-group is sufficiently large, the path via states where every M-player chooses the same action needs fewer mutations. By Lemma 3, we know that the transitions via this path require the fewest mutations under an M-player influence structure. Under an M-player influence structure, the transition cost from the risk-dominant equilibrium to the payoff dominant equilibrium is independent of the population size of both groups. While the transition from the payoff dominant equilibrium to the risk-dominant equilibrium needs a fraction of players in F-group to make mistakes. The risk-dominant equilibrium is selected when the population size of the F-group is small, and the payoff dominant equilibrium will be selected as the population size of the F-group gets larger. In particular, when the population size of the F-group is sufficiently large ($f \geqslant \ell + \frac{(a-d)\ell}{b-d}$), under an M-player influence structure, the mutation of one M-player is sufficient to reach the payoff dominant equilibrium. Thus, $\overrightarrow{B}_g$ will be the LRE in this case.

5 Conclusion

We have presented a model of social coordination and network formation between two groups of players, where the players of one group support a limited number of links to the players of the other group. This paper has shown that the population sizes of the two groups have a powerful impact on the equilibrium selection. When the both populations are large in comparison to the number of links that may be sustained, the players of two groups will coordinate on the payoff dominant action. On the contrary, if either of the two groups is small, all players will choose the risk-dominant equilibrium.

There are several natural extensions to the research presented here. First, it would be desirable to consider the case where the cost of supporting and maintaining a link is high. As shown in Goyal and Vega-Redondo (2005) and Staudigl and Weidenholzer (2014), for high costs of supporting links, we could find that payoff dominant convention arises for a wider range of parameters. Second, one could imagine a model where the F-players may reject or accept incoming links. In this case, we could find the states with mixed strategies become absorbing.

References

Alós-Ferrer, C. & Weidenholzer, S. (2008), 'Contagion and efficiency', *Journal of Economic Theory* 143(1), 251–274.

Anwar, A. W. (2002), 'On the co-existence of conventions', *Journal of Economic Theory* 107(1), 145–155.

Bhaskar, V. & Vega-Redondo, F. (2004), 'Migration and the evolution of conventions', *Journal of Economic Behavior & Organization* 55(3), 397–418.

Bilancini, E., Boncinelli, L. et al. (2014), Social coordination with locally observable types, Technical report, University of Modena and Reggio E., Dept. of Economics" Marco Biagi".

Blume, L. E. (1993), 'The statistical mechanics of strategic interaction', *Games and Economic Behavior* 5(3), 387–424.

Blume, L. E. (1995), 'The statistical mechanics of best-response strategy revision', *Games and Economic Behavior* 11(2), 111–145.

Ellison, G. (1993), 'Learning, local interaction, and coordination', *Econometrica: Journal of the Econometric Society* pp. 1047–1071.

Ellison, G. (2000), 'Basins of attraction, long-run stochastic stability, and the speed of step-bystep evolution', The *Review of Economic Studies* 67(1), 17–45.

Ely, J. C. (2002), 'Local conventions', *Advances in Theoretical Economics* 2(1).

Gale, D. & Shapley, L. S. (1962), 'College admissions and the stability of marriage', *American mathematical monthly* pp. 9–15.

Goyal, S. & Vega-Redondo, F. (2005), 'Network formation and social coordination', *Games and Economic Behavior* 50(2), 178–207.

Harsanyi, J. C. & Selten, R. (1988), 'A general theory of equilibrium selection in games', MIT Press Books 1.

Hojman, D. A. & Szeidl, A. (2006), 'Endogenous networks, social games, and evolution', *Games and Economic Behavior* 55(1), 112–130.

Jackson, M. O. & Watts, A. (2002), 'On the formation of interaction networks in social coordination games', *Games and Economic Behavior* 41(2), 265–291.

Jackson, M. O. & Watts, A. (2010), 'Social games: Matching and the play of finitely repeated games', *Games and Economic Behavior* 70(1), 170–191.

Jackson, M. & Watts, A. (2008), 'Equilibrium Existence in Bipartite Social Games: A Generalization of Stable Matchings', *Economics Bulletin* 3(12), 1–8.

Kandori, M., Mailath, G. J. & Rob, R. (1993), 'Learning, mutation, and long run equilibria in games', *Econometrica: Journal of the Econometric Society* pp. 29–56.

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

Kranton, R. E. & Minehart, D. F. (2003), A theory of buyer-seller networks, in 'Networks and Groups', Springer, pp. 347–378.

Oechssler, J. (1997), 'Decentralization and the coordination problem', *Journal of Economic Behavior & Organization* 32(1), 119–135.

Pongou, R. & Serrano, R. (2013), 'Fidelity networks and long-run trends in hiv/aids gender gaps', The *American Economic Review* 103(3), 298–302.

Roth, A. E. & Sotomayor, M. (1989), 'The college admissions problem revisited', *Econometrica: Journal of the Econometric Society* pp. 559–570.

Staudigl, M. & Weidenholzer, S. (2014), 'Constrained interactions and social coordination', *Journal of Economic Theory* 152, 41–63.

Weidenholzer, S. (2010), ‘Coordination games and local interactions: a survey of the game theoretic literature’, *Games* 1(4), 551–585.

Young, H. P. (1993), ‘The evolution of conventions’, *Econometrica: Journal of the Econometric Society* pp. 57–84.

Appendix

Proof of Lemma 3: First, we consider the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states where $a^M = \overrightarrow{B}$. Thus, we start with the case where the mutations only occur among the F-players. Let m^{AB} be the minimal number of F-players switching from A to B, such that every M-player will choose B when given revision opportunity. According to Table 1, whenever the number of B-players in the F-group is larger or equal to ℓ, the M-players will choose B. They will also prefer B over A when the M-players choosing B cannot fill up all their links to B-players ($f - f^A < \ell$) in the following two cases: in the case where at the relevant threshold the M-players choosing A fill up all of their links ($f^A \geqslant \ell$), every M-player will choose B if there are less than $(f - \psi_1)$ F-players choosing A; and in the case where M-players choosing A cannot fill up all of their links ($f^A < \ell$), the M-players will prefer B if there are less than $(f - \psi_2)$ F-players choosing A. Thus, we consider the following two cases:

i) When $f^A \geqslant \ell$ after the mutations, it follows from Table 1 that every M-player will choose B with positive probability if no less than ψ_1 F-players are choosing B. Thus, in this case $m^{AB} = \lceil \psi_1 \rceil$. This case will happen if the remaining number of A-players in the F-group is larger or equal to ℓ, that is $f - \lceil \psi_1 \rceil \geqslant \ell$. Since $f, \ell \in \mathbb{N}$, this holds when $f - \psi_1 \geqslant \ell$, which yields $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$.

ii) When $f^A < \ell$ after the mutations, it follows from Table 1 that every M-player will choose B with positive probability if there are less than $(f - \psi_2)$ F-players choosing A, hence, at least ψ_2 F-players should switch to B. Thus, in this case $m^{AB} = \lceil \psi_2 \rceil$. This case will happen if the remaining number of A-players in the

F-group is smaller than ℓ, that is $f - \lceil \psi_2 \rceil < \ell$, which holds if $f - \psi_2 < \ell$. This translates into $f < \ell + \frac{(a-d)\ell}{b-d}$.

Recall that in addition we need m^{AB} to be less than ℓ. We find that if $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, it is always true that $\psi_1 < \ell$. When $f < \ell + \frac{(a-d)\ell}{b-d}$, we need $\psi_2 < \ell$, which translates into $f < \ell + \frac{b-c}{a-c}\ell$. Since $\ell + \frac{b-c}{a-c}\ell > 2\ell$, $\psi_2 < \ell$ always holds when $f < \ell + \frac{(a-d)\ell}{b-d}$. Thus, we find that:

$$m^{AB} = \begin{cases} \lceil \psi_1 \rceil & \text{if } f \geqslant \ell + \frac{(a-d)\ell}{b-d} \\ \lceil \psi_2 \rceil & \text{if } \ell < f < \ell + \frac{(a-d)\ell}{b-d} \end{cases} .$$

Recall that F-players with incoming links choose their action based on the action distribution in their neighborhood. Thus, if an F-player has only one incoming link, her action choice only depends on the action of her sole opponent. This observation will influence the nature of the transition among the absorbing sets. Then, in particular, consider a set of a set of F-players, such that each of these F-players has only one incoming link from the same M-player, if this M-player switches, all F-players link to him switch to the same action as he chooses. Now we want to understand under which interaction structure, one M-player can influence the highest number of F-players. In any absorbing state, M-players will form all of their ℓ links. Thus, $(m-1)$ M-players can have all their links to a subset of ℓ F-players. This leaves $(f-\ell)$ F-players for the remaining M-player to connect to. Since this M-player will support ℓ links, the number of F-players that only connect to this M-player is given by $\min\{\ell, f-\ell\}$. Figure 2(a) illustrates an absorbing state where two F-players only have one incoming link from one M-player. With one mutation of the M-player, there will be $\min\{\ell, f-\ell\}$ F-players switching with positive probability. We name such an interaction structure the M-player inference structure.

Now we consider the transition cost from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states where $a^M = \overrightarrow{B}$ in the aforementioned interaction structure. We already know that in an M-player inference structure, the number of F-players with one incoming link depends on the number of F-players (f), thus, we consider two cases when $f \geqslant 2\ell$ and when $2\ell > f > \ell$.

First, if $f \geqslant 2\ell$, in an M-player inference structure, one M-player can support all her ℓ links to the F-players, each of whom has the only one incoming link from the M-player. Thus, if the M-player switches to B, there will be ℓ F-players choosing B with positive probability. We know that when $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, we need $m^{AB} = \lceil \psi_1 \rceil$ F-players to choose B so that B is the best reply for every M-player. Since it is always true that $\ell > \psi_1$, when $f \geqslant 2\ell$, we are able to reach a state in which $a^M = \overrightarrow{B}$ with just one mutation, that is $P_M(A, B) = 1$.

Second, if $2\ell > f > \ell$, in an M-player inference structure, one M-player can link to $(f - \ell)$ F-players, each of whom has the only incoming link from the M-player. We know that the transition from $\overrightarrow{A}_g$ to a state where $a^M = \overrightarrow{B}$ requires at least m^{AB} F-players to switch from A to B, so we distinguish two cases: i) when $2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d}$ and ii) when $\ell + \frac{(a-d)\ell}{b-d} > f > \ell$. i) When $2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, we need $\lceil \psi_1 \rceil$ F-players to choose B so that every M-player will prefer B over A. Consider an M-player inference structure (as Figure 2(a)), in which one M-player connects to the highest number of F-players, each of whom has the only incoming link from this M-player. If this M-player switches to B, $(f - \ell)$ F-players will switch to B with positive probability. Since it is always true that $f - \ell \geqslant \lceil \psi_1 \rceil$ in this case, we are able to reach a state in which $a^M = \overrightarrow{B}$ with just one mutation, that is $P_M(A, B) = 1$. ii) When $\ell + \frac{(a-d)\ell}{b-d} > f > \ell$, we need $\lceil \psi_2 \rceil$ F-players to choose B so that every M-player will prefer B over A. Given an M-player inference structure, if one M-player switches to B, $(f - \ell)$ F-players will switch to B with positive probability. Since $f - \ell < \lceil \psi_2 \rceil$ in this case, in addition to the M-player, we still need that $[\lceil \psi_2 \rceil - (f - \ell)]$ F-players who are not only connected to the M-player to also switch to B by mistakes. Thus, when $\ell + \frac{(a-d)\ell}{b-d} > f > \ell$, we find that $P_M(A, B) = \lceil \psi_2 \rceil - (f - \ell) + 1$. To sum up, we have:

$$P_M(A,B) = \begin{cases} 1 & \text{if } f \geqslant \ell + \frac{(a-d)\ell}{b-d} \\ \lceil \psi_2 \rceil - (f-\ell) + 1 & \text{if } f < \ell + \frac{(a-d)\ell}{b-d} \end{cases}.$$

Next, we consider the transition from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$ via states where $a^M = \overrightarrow{A}$. We denote by m^{BA} the minimal number of F-players switching from B to A, such that every M-player will choose A when given revision opportunity. According to Table 1, whenever the number of B-players in the F-group is larger or equal to ℓ, the M-players will never choose A. Thus, for M-players to choose A, $f - f^A < \ell$ must be true. Now there are two cases where A is the best reply for every M-player. i) When the M-players choosing A can fill up all of their links to the A-players in F-group ($f^A \geqslant \ell$), we can infer from Table 1 that $m^{BA} = \lceil f - \psi_1 \rceil$ in this case. ii) When the M-players choosing A cannot fill up all of their links to the A-players in F-group ($f^A < \ell$), according to Table 1, we have $m^{BA} = \lceil f - \psi_2 \rceil$ in this case. The first case happens if the number of F-players choosing A is larger or equal to ℓ, that is $\lceil f - \psi_1 \rceil \geqslant \ell$, which holds if $f - \psi_1 \geqslant \ell$, hence we have $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$. The second case happens if the number of F-players choosing A is less than ℓ, that is $\lceil f - \psi_2 \rceil < \ell$, which holds if $f - \psi_2 < \ell$. We can translate this into $f < \ell + \frac{(a-d)\ell}{b-d}$. Recall that in both cases, the remaining number of F-players choosing B should be less than ℓ, that is $f - m^{BA} < \ell$. When $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, it must be the case that $f - m^{BA} < \ell$. When $f < \ell + \frac{(a-d)\ell}{b-d}$, we need that $f - \lceil f - \psi_2 \rceil < \ell$. From the analysis above, we know that $\psi_2 < \ell$ must be true, thus, we always have $f - m^{BA} < \ell$ when $f < \ell + \frac{(a-d)\ell}{b-d}$. Thus, we have:

$$m^{BA} = \begin{cases} \lceil f - \psi_1 \rceil & \text{if } f \geqslant \ell + \frac{(a-d)\ell}{b-d} \\ \lceil f - \psi_2 \rceil & \text{if } f < \ell + \frac{(a-d)\ell}{b-d} \end{cases}.$$

To complete the transition from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$ via states where $a^M = \overrightarrow{A}$ with the minimal number of mutations, again we consider the M-player inference structures. In such an interaction structure, if one M-player switches to A, $\min\{\ell, f-\ell\}$ F-players will choose A with positive probability. We distinguish two cases when $f \geqslant 2\ell$ and when $2\ell > f > \ell$.

First, if $f \geqslant 2\ell$, in an M-player inference structure, one M-player can support all her ℓ links to the F-players, each of whom has the only incoming link from the M-player. We know that when $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, we need $m^{AB} = \lceil f - \psi_1 \rceil$ F-players to choose A so that A is the best reply for every M-player. Once the M-player switches to A by mistakes, there will be ℓ F-players choosing A with positive probability. Since it must be the case that $\lceil f - \psi_1 \rceil > \ell$ when $f \geqslant 2\ell$, to make A the best reply for every M-player, in addition to the M-player, we need $(\lceil f - \psi_1 \rceil - \ell)$ F-players who are not connected to the M-player to switch to A. Thus, we find that $P_M(B, A) = \lceil f - \psi_1 \rceil - \ell + 1$ in this case.

Second, if $2\ell > f > \ell$, in an M-player inference structure, one M-player can link to $(f - \ell)$ F-players, each of whom has only one incoming link from the M-player. We know that the transition from $\overrightarrow{B}_g$ to a state where $a^M = \overrightarrow{A}$ requires at least m^{BA} F-players to choose A, so we distinguish two cases when $2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d}$ and when $\ell + \frac{(a-d)\ell}{b-d} > f > \ell$. i) When $2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, every M-player will prefer A over B if there are $\lceil f - \psi_1 \rceil$ F-players choosing A. Once the M-player switches to A by mistakes, there will be $(f - \ell)$ F-players choosing A with positive probability. Since it must be the case that $\lceil f - \psi_1 \rceil > f - \ell$ when $2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, to make A the best reply for every M-player, in addition to the M-player, we need $[\lceil f - \psi_1 \rceil - (f - \ell)]$ F-players who are not only connected to the M-player to switch to A. So we have $P_M(B, A) = \lceil f - \psi_1 \rceil - (f - \ell) + 1$ in this case. ii) When $\ell + \frac{(a-d)\ell}{b-d} > f > \ell$, every M-player will prefer A over B if there are $\lceil f - \psi_2 \rceil$ F-players choosing A. Once the M-player switches to A by mistakes, there will be $(f - \ell)$ F-players choosing A with positive probability. Since it must be the case that $\lceil f - \psi_2 \rceil > f - \ell$ when $\ell + \frac{(a-d)\ell}{b-d} > f > \ell$, in addition to the M-player, we need $[\lceil f - \psi_2 \rceil - (f - \ell)]$ F-players who are not only connected to the M-player to switch to A, so that every M-player prefer A over B. Thus, we have $P_M(B, A) = \lceil f - \psi_2 \rceil - (f - \ell) + 1$ in this case. To sum up, we have:

$$P_M(B,A) = \begin{cases} \lceil f-\psi_1 \rceil - \ell + 1 & \text{if } f \geqslant 2\ell \\ \lceil f-\psi_1 \rceil - (f-\ell) + 1 & \text{if } 2\ell > f \geqslant \ell + \frac{(a-d)\ell}{b-d} \\ \lceil f-\psi_2 \rceil - (f-\ell) + 1 & \text{if } \ell < f < \ell + \frac{(a-d)\ell}{b-d} \end{cases}.$$

Proof of Lemma 4: First, we consider the transition from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states where $a^F = \overrightarrow{B}$. Thus, we consider the case where the mutations occur among the M-players. We denote by f^{AB} the minimal number of M-players switching from A to B, such that every F-player will choose B when given revision opportunity. We claim that in this case, $f^{AB} \geqslant \lceil m(1-q^*) \rceil$, and prove it via contradiction. Assume that $f^{AB} = \gamma$ and $\gamma < \lceil m(1-q^*) \rceil$, that is when γ M-players switch to B by mistakes, every F-player will choose B with positive probability. Recall that if an F-player i has at least one incoming link ($\sum_{j\in M} g_{ji} > 0$), she will switch to B when at least $(1-q^*)$ of her neighbors choose B; if she has no incoming link ($\sum_{j\in M} g_{ji} = 0$), she will switch to B when at least $(1-q^*)$ of M-players choose B. Since $\gamma < \lceil m(1-q^*) \rceil$, an F-player with no incoming link will not switch to B when $f^{AB} = \gamma$. Thus, we consider an interaction structure where every F-player has at least one incoming link. Under this interaction structure, the players from M-group support $m\ell$ links in all, and there are $\gamma\ell$ links supported by B-players. To make sure that every F-player prefers B over A, we need that for every F-player, at least $\lceil (1-q^*) \rceil$ of her neighbors choose B, that is we need $d_i^B \geqslant \lceil d_i(1-q^*) \rceil$ for $\forall i \in F$. Summing up all the F-players, we need that $\sum_{i\in F} d_i^B \geqslant \sum_{i\in F} \lceil d_i(1-q^*) \rceil$. For the right hand side of the inequality, according to the property of ceiling function, $\lceil x \rceil + \lceil y \rceil \geqslant \lceil x+y \rceil$, we have $\sum_{i\in F} \lceil d_i(1-q^*) \rceil \geqslant \lceil (1-q^*) \sum_{i\in F} d_i \rceil = \lceil (1-q^*) m\ell \rceil$. We also know that the left hand side of the inequality represents the total number of links supported by B-players, so we have $\sum_{i\in F} d_i^B = \gamma\ell$. Thus, the inequality implies that $\gamma\ell \geqslant \lceil m\ell(1-q^*) \rceil$, which contradicts $\gamma < \lceil m(1-q^*) \rceil$. So we prove the claim that $f^{AB} \geqslant \lceil m(1-q^*) \rceil$. Thus, to reach a state where $a^F = \overrightarrow{B}$ with the minimal number of mutations, we consider the following interaction structure: All the M-players support their ℓ links to the same set of F-players, and the rest of the F-players are left unlinked (See Figure 2(b) for instance). We name such an interaction structure the F-player inference structure. Given this interaction structure, every F-player with ℓ incoming links will choose B with positive probability if at least $\lceil m(1-q^*)\rceil$ M-players switch to B by mutation. Once these mutations occur, the F-players with no incoming links ($\sum_{j\in M} g_{ji} = 0$) will also choose B as the best reply to the action distribution in M-group. Thus, the minimal transition cost from $\overrightarrow{A}_g$ to $\overrightarrow{B}_g$ via states in which $a^F = \overrightarrow{B}$ is $\lceil m(1-q^*)\rceil$, establishing that $P_F(A, B) = \lceil m(1-q^*)\rceil$.

Next, we consider the transition from $\overrightarrow{B}_g$ to $\overrightarrow{A}_g$ via states where $a^F = \overrightarrow{A}$. To reach a state where $a^F = \overrightarrow{A}$ with the minimal number of mutations, we still consider the F-player inference structure as Figure 2(b). Analogously to the previous analysis, every F-player will choose A if at least $\lceil mq^*\rceil$ M-players switch to A. Thus, we can obtain that $P_F(B, A) = \lceil mq^*\rceil$. $\square$

Proof of Proposition 1: We already know that $CR(\overrightarrow{B}_g) = R(\overrightarrow{A}_g) = \lceil \min\{m, f\}(1-q^*)\rceil$ and $CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = \lceil \min\{m, f\}q^*\rceil$ when $f = \ell$. By risk-dominance, we know that $q^* < \frac{1}{2}$, so we have $\min\{m, f\}q^* < \min\{m, f\}(1-q^*)$, which implies that $\lceil \min\{m, f\}q^*\rceil \leqslant \lceil \min\{m, f\}(1-q^*)\rceil$. Thus, according to Lemma 1, when $f = \ell$, the sets of LRE of the model are given by $\mathcal{S} = \overrightarrow{A}_g$ if $\lceil \min\{m, f\}q^*\rceil < \lceil \min\{m, f\}(1-q^*)\rceil$; $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ if $\lceil \min\{m, f\}q^*\rceil = \lceil \min\{m, f\}(1-q^*)\rceil$. $\square$

Proof of Proposition 2: We know by Lemma 3 and Lemma 4 that $P_F(A, B) = \lceil m(1-q^*)\rceil$, $P_M(A, B) = \lceil \psi_2\rceil - (f-\ell) + 1$, $P_F(B, A) = \lceil mq^*\rceil$, and $P_M(B, A) = \lceil f - \psi_2\rceil - (f-\ell) + 1$.

(1) When $\lceil m(1-q^*)\rceil < \lceil \psi_2\rceil - (f-\ell) + 1$ and $\lceil mq^*\rceil < \lceil f - \psi_2\rceil - (f-\ell) + 1$, which implies that $m < \min\{\frac{\psi_2 - f + \ell + 1}{1-q^*}, \frac{\ell - \psi_2 + 1}{q^*}\}$, the path via states where every F-player chooses the same action requires fewer mistakes to complete the transition. Thus, when $m < \min\{\frac{\psi_2 - f + \ell + 1}{1-q^*}, \frac{\ell - \psi_2 + 1}{q^*}\}$, the radius and coradius of the absorbing

sets are given by: $CR(\overrightarrow{B}_g) = R(\overrightarrow{A}_g) = P_F(A,B) = \lceil m(1-q^*)\rceil$ and $CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = P_F(B,A) = \lceil mq^*\rceil$. By risk-dominance, we have $q^* < \frac{1}{2}$, which implies that $\lceil mq^*\rceil \leqslant \lceil m(1-q^*)\rceil$. Thus, the set of LRE are $\mathcal{S} = \overrightarrow{A}_g$ if $\lceil mq^*\rceil < \lceil m(1-q^*)\rceil$; $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ if $\lceil mq^*\rceil = \lceil m(1-q^*)\rceil$.

(2) When $\lceil m(1-q^*)\rceil \geqslant \lceil \psi_2\rceil - (f-\ell) + 1$ and $\lceil mq^*\rceil \geqslant \lceil f-\psi_2\rceil - (f-\ell)+1$, we have $m > \max\{\frac{\psi_2 - f+\ell+1}{1-q^*}, \frac{\ell-\psi_2+1}{q^*}\}$. Thus, when $m > \max\{\frac{\psi_2 - f+\ell+1}{1-q^*}, \frac{\ell-\psi_2+1}{q^*}\}$, it requires fewer mistakes to complete the transitions among absorbing sets via states where every M-player chooses the same action. Thus, the radius and coradius of the absorbing sets are given by: $CR(\overrightarrow{B}_g) = R(\overrightarrow{A}_g) = \lceil \psi_2\rceil - (f-\ell)+1$ and $R(\overrightarrow{B}_g) = CR(\overrightarrow{A}_g) = \lceil f-\psi_2\rceil - (f-\ell)+1$.

First, we find that when $\lceil \psi_2\rceil > \lceil f-\psi_2\rceil$, we have $R(\overrightarrow{A}_g) > CR(\overrightarrow{A}_g)$, establishing that $\mathcal{S} = \overrightarrow{A}_g$ in this case. If $\psi_2 \in \mathbb{Z}$, the condition holds if $\psi_2 > f-\psi_2$, which translates into $f < \frac{2(c-d)\ell}{b+c-a-d}$. If $\psi_2 \notin \mathbb{Z}$, the condition holds if $\lceil \psi_2\rceil > \frac{1}{2}(f+1)$. If f is even, we can translate the condition into $\psi_2 > \frac{1}{2}f$, which yields $f < \frac{2(c-d)\ell}{b+c-a-d}$. If f is odd, the condition implies that $\psi_2 > \frac{1}{2}(f+1)$, which yields $f < \frac{2(c-d)\ell-(a+b-c-d)}{b+c-a-d}$.

Next, we find that when $\lceil \psi_2\rceil < \lceil f-\psi_2\rceil$, we have $R(\overrightarrow{B}_g) > CR(\overrightarrow{B}_g)$, establishing that $\mathcal{S} = \overrightarrow{B}_g$ in this case. If $\psi_2 \in \mathbb{Z}$, the condition holds if $\psi_2 < f-\psi_2$, which translates into $f > \frac{2(c-d)\ell}{b+c-a-d}$. If $\psi_2 \notin \mathbb{Z}$, the condition holds if $\lceil \psi_2\rceil < \frac{1}{2}(f+1)$. If f is even, we can translate the condition into $\psi_2 < \frac{1}{2}f$, which yields $f > \frac{2(c-d)\ell}{b+c-a-d}$. If f is odd, the condition implies that $\psi_2 \leqslant \frac{1}{2}(f-1)$, which yields $f \geqslant \frac{2(c-d)\ell+(a+b-c-d)}{b+c-a-d}$.

Last, we find when $\lceil \psi_2\rceil = \lceil f-\psi_2\rceil$, we have $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$. If $\psi_2 \in \mathbb{Z}$, the condition holds when $\psi_2 = \frac{1}{2}f$, which yields $f = \frac{2(c-d)\ell}{b+c-a-d}$. If $\psi_2 \notin \mathbb{Z}$, the condition holds only if f is odd and $\lceil \psi_2\rceil = \frac{1}{2}(f+1)$, we can translate the equation into $\frac{2(c-d)\ell+(a+b-c-d)}{b+c-a-d} > f \geqslant \frac{2(c-d)\ell-(a+b-c-d)}{b+c-a-d}$.$\square$

Proof of Proposition 3: We already know by Lemma 3 and Lemma 4 that $P_F(A,B) = \lceil m(1-q^*)\rceil$, $P_M(A,B) = 1$, $P_F(B,A) = \lceil mq^*\rceil$ and $P_M(B,A) = \lceil f-\psi_1\rceil - \min\{f-\ell,\ell\}+1$. Since the transition cost between absorbing sets is a natural number and should be no less than one, thus, it must be the case that $\min\{P_M(B,A), P_F(B,A)\} \geqslant 1$ and $\lceil m(1-q^*)\rceil \geqslant 1$. So we have

$$CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = \min\{P_M(B,A), P_F(B,A)\} = 1$$

and

$$CR(\overrightarrow{A}_g) = R(\overrightarrow{B}_g) = \min\{P_M(B,A), P_F(B,A)\} \geqslant 1.$$

Since $P_M(B,A) = \lceil f - \psi_1 \rceil - \min\{f-\ell, \ell\} + 1 > 1$ must be true when $f \geqslant \ell + \frac{(a-d)\ell}{b-d}$, both absorbing sets are LRE only if $\lceil mq^* \rceil = 1$, which implies that $1 \leqslant m \leqslant \frac{1}{q^*}$. $\overrightarrow{B}_g$ is the unique LRE if $\lceil mq^* \rceil > 1$, which implies that $m > \frac{1}{q^*}$. Thus, we have $\mathcal{S} = \overrightarrow{B}_g$ if $\lceil mq^* \rceil > 1$, and $\mathcal{S} = \overrightarrow{A}_g \cup \overrightarrow{B}_g$ if $m \leqslant \frac{1}{q^*}$. $\square$