

Rapid evolution under heterogeneous linking costs

Zhiwei Cui

Renmin University of China

Jinhua Liu

liujh2917@163.com

Renmin University of China

Research Article

Keywords: Coordination game, Network formation, Heterogeneous linking costs, Learning dynamics, Waiting time

Posted Date: September 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10688020/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Rapid evolution under heterogeneous linking costs^{*}

Zhiwei Cui[†] Jinhua Liu[‡]

Abstract

We consider a model of social coordination with endogenous network formation in which agents simultaneously choose an action in a coordination game and pay heterogeneous costs to form links. When some agents incur high linking costs while others face low costs, action-heterogeneous Nash equilibria may arise. If each agent can link to fewer than half of the population, the payoff-dominant strategy profiles are stochastically stable. Crucially, the presence of high-linking-cost agents can accelerate convergence to the payoff-dominant convention. Our findings indicate that heterogeneity in linking costs shapes both equilibrium diversity and the speed of convergence toward efficient conventions.

Keywords: Coordination game, Network formation, Heterogeneous linking costs, Learning dynamics, Waiting time

JEL: C72, C73, D83, D85

^{*}This work was financially supported by the National Natural Science Foundation of China (No. 72471228) and the Fund for Building World-Class Universities (Disciplines) of Renmin University of China.

[†]School of Economics, Renmin University of China, Beijing 100872, P. R. China. Email: zhiweicuisx@gmail.com

[‡]Corresponding author. School of Mathematics, Renmin University of China, Beijing 100872, P. R. China. Email: liujh2917@163.com

1 Introduction

Coordination problems frequently arise in social and economic interactions, such as the choice of programming languages (e.g., Python versus C++) and technology standards (e.g., Blu-ray versus DVD). Usually two issues emerge in such situations: first, whether to coordinate on the same action; and, second, whether to coordinate on the most efficient action. Here we focus on the second problem by considering a model that integrates social coordination with network formation. We explore the role of heterogeneous linking costs, under which some agents incur high linking costs while others incur low ones.

Specifically, we consider a finite population of agents. Each agent forms costly links with others in a noncooperative manner, as in Bala and Goyal (2000). In particular, some agents incur a high cost per link, while others incur a low one. Agents face an exogenous constraint on the number of links that they can form, thereby capturing natural limits imposed by time, resources, and social capacity. In addition to the linking decision, each agent chooses an action in a symmetric 2×2 coordination game and uses that action uniformly in all of her interactions. A directed link from agent i to agent j enables i to interact with j , and only i receives a payoff from the interaction. The 2×2 game embodies the classic tension between efficiency and risk: one strict Nash equilibrium is payoff-dominant, while the other is risk-dominant.

High linking costs impose an additional constraint on agents' linking decisions. To fix ideas, consider a profile in which agents with high linking costs play the payoff-dominant action while agents with low linking costs play the risk-dominant action. Agents with high linking costs may find it optimal to form links only with others who play the same action, since their linking costs are so high that links to agents playing the risk-dominant action are not beneficial. In other words, these agents may not be motivated to use all of their links when they play the payoff-dominant action. In fact, these action-heterogeneous strategy profiles are Nash equilibria under certain parameter conditions (see Lemma 1 and Proposition 1). Notably, the number of agents with high linking costs must be moderate: if it is too small, high-linking-cost agents will be motivated to switch to the risk-dominant action, whereas if it is too large, low-linking-cost agents will be motivated to switch to the payoff-dominant

action. It is worth mentioning that when all agents have the same linking cost, as in Staudigl and Weidenholzer (2014) and Cui and Shi (2022), action-heterogeneous Nash equilibria do not exist.

Then we consider a perturbed myopic best-response dynamics in discrete time. Payoff-dominant profiles are stochastically stable provided that agents can form links with fewer than half of the other agents (see Theorem 1). The underlying intuition is the same as in Staudigl and Weidenholzer (2014) and Cui and Shi (2022). If about half the population plays the payoff-dominant action, it is optimal for everyone to play it as well. This implies that it is harder to leave the basin of attraction of the payoff-dominant convention than to enter it from any other absorbing set. Consequently, the payoff-dominant convention is most frequently observed in the long run.

Finally, and perhaps most importantly, we examine the expected waiting time required to enter the basin of attraction of stochastically stable states. We show that the presence of agents with high linking costs substantially lowers the order of the expected waiting time relative to the homogeneous low-linking-cost benchmark analyzed in Staudigl and Weidenholzer (2014) (see Theorem 2). The intuition is as follows. High linking costs impose additional restrictions on agents' linking decisions, making it easier to induce these agents to switch from the risk-dominant action to the payoff-dominant action. Accordingly, high-linking-cost agents serve as a catalytic minority facilitating the transition from the risk-dominant convention to the payoff-dominant convention. Put differently, a reforming effect arises: the presence of agents with high linking costs speeds up convergence to the payoff-dominant convention.

1.1 Related literature

Our paper contributes to the literature on stochastic learning in coordination games. When interaction structures are exogenously given, risk-dominant conventions are usually stochastically stable under myopic best-response updating (see, e.g., Blume 1993, Ellison 1993, Kandori, Mailath, and Rob 1993, Young 1993, Ellison 2000, and Morris 2000).¹ However,

¹When agents use imitative rules to update their strategies, payoff-dominant conventions may be selected in the long run (see, e.g., Alós-Ferrer and Weidenholzer 2008, Alós-Ferrer and Weidenholzer 2014, Cui 2014,

when interaction structures are endogenously formed, payoff-dominant conventions may be the only stochastically stable outcomes.

In particular, one strand of the literature posits that interaction structures are formed strategically by agents through a process of network formation. In Goyal and Vega-Redondo (2005), agents build links in a noncooperative and unconstrained manner and play a 2×2 coordination game with directly connected agents. Unconstrained network formation invariably gives rise to complete interaction networks.² Therefore, only the risk-dominant convention emerges as stochastically stable, which is consistent with the case of fixed interaction structures. In contrast to Goyal and Vega-Redondo (2005), Staudigl and Weidenholzer (2014) assume that agents can form only a limited number of links and show that only the payoff-dominant convention is selected in the long run, provided that payoffs accrue solely to link initiators. Cui and Shi (2022) argue that the main result in Staudigl and Weidenholzer (2014) extends to a general class of size-dependent coordination games. Building on Staudigl and Weidenholzer (2014), Cui (2023) explores the role of linking frictions in accelerating convergence to the payoff-dominant convention. In Bilancini and Boncinelli (2018), agents are categorized into two types, and actions may serve as signals of these types, enabling coordination across them. Furthermore, Lu and Shi (2023) consider size-dependent minimum-effort games where agents face different constraints on link formation and demonstrate that the payoff-dominant convention still prevails in the long run when enough agents face tight constraints. In addition, given the linking constraint, Cui and Weidenholzer (2021) assume that agents receive payoffs both from interactions with agents who form links to them and from interactions with agents to whom they form links, and show that different agents may choose different actions in the long run. As a complement, the present paper studies the role of heterogeneous linking costs and shows that the presence of a small group of agents with high linking costs can accelerate convergence to the payoff-dominant convention.

The role of heterogeneity has also received attention in the broader literature on stochastic learning in coordination games. In Neary (2012), agents have heterogeneous preferences and may choose their most preferred action in the long run. Alós-Ferrer and Shi (2012) investigate

and Cui and Wang (2016).

²This mechanism also underlies the selection of the risk-dominant convention in Jackson and Watts (2002), in which bilateral mutual consent is mandatory for link formation.

the effect of heterogeneous memory lengths on the long-run outcomes of imitation dynamics and show that such heterogeneity is crucial for equilibrium selection. Newton (2021) shows that the asymmetry property of the stochastic learning dynamics proposed by Peski (2010) can be defined at the level of the behavioral rule, and then examines three dimensions of heterogeneity: within individual agents, across different agents, and in the timing of strategy revision. In a bilingual game, Naono (2022) shows that heterogeneous costs of playing a bilingual strategy can lead to its long-run persistence. Unlike these contributions, the present work studies the effect of heterogeneous linking costs on long-run outcomes, assuming agents can strategically choose their interaction partners.

Interestingly, one strand of the literature on stochastic learning studies the expected waiting time for convergence to stochastically stable equilibria. To the best of our knowledge, Ellison (1993) was the first to study this waiting time. Ellison (2000) develops two new measures—the coradius and the modified coradius—which provide upper bounds on the order of the expected waiting time. Furthermore, Cui and Zhai (2010) introduce iterative cycle decompositions from Freidlin and Wentzell (1984) to provide an exact measure of this order.³ As an application of Ellison (2000), Norman (2009) shows that inertia driven by switching costs can lead to more rapid evolution in a class of games. Newton and Angus (2015) point out that local coalitional behavior can both slow and hasten convergence to the most efficient convention. Cui (2023) demonstrates that moderate linking frictions can accelerate convergence to the payoff-dominant convention. The present work contributes to this literature by showing that heterogeneous linking costs can accelerate convergence to the payoff-dominant convention.

2 The Model

We consider a finite population of n agents who play a 2×2 coordination game against each other. Let $I = \{1, 2, \dots, n\}$ be the set of all agents. In addition to choosing actions within the game, agents select their interaction partners through link formation.⁴

³Newton and Sandholm (2022) show that iterative cycle decompositions are equivalent to the aggregative stage of Edmonds’ algorithm, and this equivalence can be used to recover well-known results.

⁴We follow the notation of Neary (2012) and of Staudigl and Weidenholzer (2014).

Each player chooses an action from the set $\{A, B\}$. Let $u(\alpha, \alpha') > 0$ denote the payoff of a player with action α against another player with action α' . The payoffs are given in the following table:

	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 strict Nash equilibria. We further assume that $b > a$ so that (B, B) is payoff-dominant, and that $a + c > b + d$ so that (A, A) is risk-dominant in the sense of Harsanyi and Selten (1988). In other words, action A is the unique best response when the opponent plays both actions with equal probability. By $b > a$ and $a + c > b + d$, we have $c > d$. Finally, we assume $a > c$, implying that an A -player prefers playing against another A -player over playing against a B -player. These assumptions imply the following ordering of payoffs:

$$b > a > c > d > 0.$$

In addition to their action choice in the 2×2 coordination game, agents decide with whom to link. If agent i forms a link with agent j , we write $g_{ij} = 1$; otherwise, we write $g_{ij} = 0$. We assume that agents cannot link to themselves, i.e., $g_{ii} = 0$ for all $i \in I$. We also assume that each agent may support at most $k \leq n - 1$ links, i.e., $\sum_{j \in I} g_{ij} \leq k$ for all $i \in I$. However, agents are not constrained by the number of links they can passively receive.

The linking decision of agent i can be summarized by an n -tuple $g_i = (g_{i1}, g_{i2}, \dots, g_{in}) \in \mathcal{G}_i$, where $\mathcal{G}_i = \{g_i \in \{0, 1\}^n : g_{ii} = 0 \text{ and } \sum_{j \in I} g_{ij} \leq k\}$ is the set of all permissible linking decisions of agent i . We denote by $g = (I, (g_i)_{i \in I}) \in \mathcal{G}$ the network induced by the linking decisions of all agents, where $\mathcal{G} = \prod_{i \in I} \mathcal{G}_i$ is the set of all permissible networks. We denote by $N_i^{\text{out}}(g) = \{j \in I : g_{ij} = 1\}$ the set of agents to whom agent i forms a link. We refer to the agents in $N_i^{\text{out}}(g)$ as the neighbors of agent i . Finally, we denote by $d_i^{\text{out}}(g) = \sum_{j \in I} g_{ij}$ the out-degree of agent i , i.e., the number of links actively formed by agent i .

A pure strategy of agent $i \in I$ consists of her action choice α_i and her linking decision

g_i , i.e., $s_i = (\alpha_i, g_i) \in \mathcal{S}_i = \{A, B\} \times \mathcal{G}_i$. A strategy profile of all agents is an n -tuple $s = (s_i)_{i \in I} \in \mathcal{S} = \prod_{i \in I} \mathcal{S}_i$. The strategy profile of all agents except i is an $(n - 1)$ -tuple $s_{-i} = (s_1, \dots, s_{i-1}, s_{i+1}, \dots, s_n) \in \mathcal{S}_{-i} = \prod_{j \in I \setminus \{i\}} \mathcal{S}_j$.

We denote by $I_A = \{i \in I | \alpha_i = A\}$ the set of A -players and by $I_B = \{j \in I | \alpha_j = B\}$ the set of B -players. Let $m = |I_A|$ denote the number of A -players. Then, the number of B -players is given by $n - m$.

Each agent i incurs a cost γ_i for maintaining each of its links. We assume that there are two types of linking costs: a low cost, γ^l , and a high cost, γ^h , where $0 \leq \gamma^l \leq d$ and $d < \gamma^h \leq c$.⁵ Let $I^h = \{i \in I | \gamma_i = \gamma^h\}$ denote the set of agents with the high linking cost γ^h . The set of agents with the low linking cost γ^l is then $I \setminus I^h$, which we denote by I^l . Let $n^h = |I^h|$ denote the number of agents with the high linking cost. Accordingly, the number of agents with the low linking cost is $n - n^h$, which we denote by n^l .

Given the profile $s \in \mathcal{S}$, the payoffs of agent i are

$$\pi_i(s) = \sum_{j \in I} g_{ij} u(\alpha_i, \alpha_j) - \gamma_i d_i^{\text{out}}(g).$$

The description of payoffs concludes the specification of the social game $\Gamma = (I, \{\mathcal{S}_i\}_{i \in I}, \{\pi_i\}_{i \in I})$.

3 Nash equilibria

There are a variety of Nash equilibria. First, strategy profiles in which every agent chooses the same action and maintains k links constitute Nash equilibria, provided that $\gamma^l < \gamma^h \leq c$. We refer to such equilibria as *action-homogeneous Nash equilibria*. Nash equilibria with heterogeneous actions across agents may also emerge. We term these *action-heterogeneous Nash equilibria*. The following lemma characterizes the action-heterogeneous Nash equilibria; its proof can be found in the Appendix.

Lemma 1. *For any action-heterogeneous Nash equilibrium, the number of B -players is strictly smaller than k . Furthermore, all the agents in I^l choose action A and all the agents*

⁵In fact, our results continue to hold if the linking costs of agents in I^l differ, provided that $0 \leq \gamma_i \leq d$ for all $i \in I^l$, and if the linking costs of agents in I^h differ, provided that $d < \gamma_i \leq c$ for all $i \in I^h$.

in I^h choose action B .

Trivially, in an action-heterogeneous Nash equilibrium, the number of agents choosing the payoff-dominant action B should be strictly less than k ; otherwise, all remaining agents have incentives to switch to action B . More importantly, all the agents with the same linking cost choose the same action. More specifically, all the agents with the low linking cost γ^l choose the risk-dominant action A and all the agents with the high linking cost γ^h choose the payoff-dominant action B . The following proposition establishes a necessary and sufficient condition for the existence of action-heterogeneous Nash equilibria. Its proof is relegated to the Appendix.

Proposition 1. *Action-homogeneous Nash equilibria always exist. Action-heterogeneous Nash equilibria exist if and only if $\frac{(a-c)\min\{n^l, k\} + (c-\gamma^h)k}{b-\gamma^h} + 1 \leq n^h \leq \frac{(a-c)\min\{n^l-1, k\} + (c-d)k}{b-d}$.*

Proposition 1 shows that when agents have different linking costs, there are action-heterogeneous Nash equilibria, provided that

$$\frac{(a-c)\min\{n^l, k\} + (c-\gamma^h)k}{b-\gamma^h} + 1 \leq n^h \leq \frac{(a-c)\min\{n^l-1, k\} + (c-d)k}{b-d}.$$

The intuition is as follows. Given that $\gamma^h > d$, agents in I^h refrain from forming links with A -players when they choose action B . This leaves agents in I^h facing tighter linking constraints under the condition $n^h < k$. Hence, fewer B -agents are needed to induce the prevalence of action B in I^h than in I^l . Furthermore, n^h cannot be excessively small relative to n^l ; otherwise, all the agents in I^h would have a strict incentive to deviate from action B to action A . Similarly, an excessively large n^h induces all the agents in I^l to switch from action A to action B .

In Staudigl and Weidenholzer (2014), agents are assumed to be homogeneous and to face identical linking costs, which leads all agents to choose the same action. Consequently, action-heterogeneous Nash equilibria cannot arise in their framework. By contrast, our model allows linking costs to vary across agents. In particular, the condition $\gamma^h > d$ for high-linking-cost agents imposes additional constraints on link formation, thereby enabling the emergence of action-heterogeneous Nash equilibria.

4 Myopic best response learning

We consider a model of perturbed myopic best-response learning à la Kandori, Mailath, and Rob (1993) and Young (1993). In each period $t = 0, 1, 2, \dots$, with a positive and independent probability $\rho \in (0, 1)$, each agent i receives an opportunity to update her strategy. When such an opportunity arises, agent i chooses a strategy to maximize her payoff in the previous period. Formally,

$$s_i(t) \in \arg \max_{s_i \in \mathcal{S}_i} \pi_i(s_i, s_{-i}(t-1)) \quad (1)$$

where if multiple strategies are suggested, agent i chooses one of them at random and each is chosen with a positive probability. With the remaining probability $1 - \rho$, agents exhibit inertia and simply maintain the status quo.

The revision process defines a Markov chain $\{S(t)\}_{t \in \mathbb{N}}$ over the set of profiles $\mathcal{S}$. We call this process the *unperturbed dynamics*. An *absorbing set* is a minimal subset $\mathcal{S}' \subset \mathcal{S}$ such that once the dynamics arrives there, the probability of leaving it is zero. Denote by $\mathcal{S}^{**}$ the set of absorbing sets.

To address the multiplicity problem, we introduce random perturbations and conduct a stochastic stability analysis (Foster and Young (1990), Kandori, Mailath, and Rob (1993), and Young (1993)). With probability $\epsilon \in (0, 1)$, an updating agent ignores the prescription of the adjustment process and chooses a strategy from $\mathcal{S}_i$ randomly according to a prespecified distribution with full support. The mistakes are assumed to be independent across agents and over time. This gives rise to *perturbed dynamics* $\{S^\epsilon(t)\}_{t \in \mathbb{N}}$. For each ϵ , $\{S^\epsilon(t)\}_{t \in \mathbb{N}}$ is an irreducible and aperiodic Markov chain and has a unique invariant distribution $\mu(\epsilon)$. It is a well-established result that $\mu^* = \lim_{\epsilon \rightarrow 0} \mu(\epsilon)$ exists and is an invariant distribution of the unperturbed dynamics. The states in the support of μ^* are called *stochastically stable states*. When the unperturbed dynamics is slightly perturbed, the dynamics will spend the vast majority of the time in such states. Denote the set of stochastically stable states by $\mathcal{S}^* = \{s \in \mathcal{S} : \mu^*(s) > 0\}$.

To explore the effect of heterogeneous linking costs on the long-run behavior of the per-

turbed dynamics, we study the expected waiting time (Ellison (1993), Ellison (2000), and Young (1998)). Let $W(s, \mathcal{S}^*, \epsilon)$ denote the expected waiting time until the set of stochastically stable states, $\mathcal{S}^*$, is first reached, given that the perturbed dynamics starts from a profile s . Then, $W(\mathcal{S}^*, \epsilon) = \max_{s \notin D(\mathcal{S}^*)} W(s, \mathcal{S}^*, \epsilon)$ can measure how quickly the perturbed dynamics converges to the set of stochastically stable states, where $D(\mathcal{S}^*)$ is the basin of attraction of $\mathcal{S}^*$. If this maximum expected waiting time is short, the convergence is fast, and $\mathcal{S}^*$ can be regarded as a good prediction for what we might expect to see even in the medium run.

4.1 Analysis of unperturbed dynamics

We denote by $\vec{A}$ and $\vec{B}$ the sets of strategy profiles where each agent chooses the same actions A and B and forms k links, respectively. Formally, for any action $\alpha \in \{A, B\}$,

$$\vec{\alpha} = \{s \in \mathcal{S} : \alpha_i = \alpha \text{ and } d_i^{\text{out}}(g) = k, \forall i \in I\}.$$

Note that an agent is indifferent toward which of the other agents to connect to. This indifference means that the transition from any element in $\vec{\alpha}$ to another element in finite periods has a positive probability. In addition, we denote by $\vec{AB}$ the set of strategy profiles where each agent in I^l chooses action A and forms k links with other agents in I^l , and each agent in I^h chooses action B and forms $n^h - 1$ links with other agents in I^h . Formally,

$$\begin{aligned} \vec{AB} = \{s \in \mathcal{S} : \alpha_i = A, d_i^{\text{out}}(g) = k, \text{ and } N_i^{\text{out}}(g) \subseteq I^l, \forall i \in I^l; \\ \alpha_j = B \text{ and } N_j^{\text{out}}(g) = I^h \setminus \{j\}, \forall j \in I^h\}. \end{aligned}$$

The following proposition characterizes the absorbing sets; its proof is provided in the Appendix.

Proposition 2. *The absorbing sets are $\vec{A}$, $\vec{B}$, and $\vec{AB}$ when $\frac{(a-c)\min\{n^l, k\} + (c-\gamma^h)k}{b-\gamma^h} + 1 < n^h < \frac{(a-c)\min\{n^l-1, k\} + (c-d)k}{b-d}$ and are $\vec{A}$ and $\vec{B}$ otherwise.*

The underlying intuition is as follows. From any initial state, the unperturbed dynamics can reach a state in which the agents of each group choose the same action with a positive

probability, by the same logic used in the proof of Lemma 1 in Staudigl and Weidenholzer (2014). Furthermore, using reasoning similar to that used in the proof of Proposition 1, the action-heterogeneous absorbing set may arise.

4.2 Analysis of perturbed dynamics

We now investigate the long-run behavior of the perturbed dynamics. To facilitate comparison with the literature (e.g., Staudigl and Weidenholzer (2014) and Cui and Shi (2022)), we focus on the case where $k \leq \frac{n-1}{2}$.

We first state the following theorem characterizing the stochastically stable states. Its proof is provided in the Appendix.

Theorem 1. *Suppose that $k \leq \frac{n-1}{2}$. Then, $\mathcal{S}^* = \vec{B}$.*

The intuition behind Theorem 1 is analogous to that underlying the main findings in Staudigl and Weidenholzer (2014) and Cui and Shi (2022). When at least $k + 1$ agents play the payoff-dominant action B , it becomes optimal for all remaining agents to choose B . Since $k \leq \frac{n-1}{2}$, exiting the basin of attraction of $\vec{B}$ requires considerably more mutations than does entering it from any other absorbing set.

To facilitate exposition, we introduce new notation. For any two functions $f(\cdot)$ and $g(\cdot)$, we write “ $f(\epsilon) = O(g(\epsilon))$ as $\epsilon \rightarrow 0$ ” as a shorthand for “there exists $C > 0$ and $\bar{\epsilon} > 0$ such that $|f(\epsilon)| < C|g(\epsilon)|$ for all $\epsilon \in (0, \bar{\epsilon})$.” Furthermore, we write “ $f(\epsilon) = \Theta(g(\epsilon))$ as $\epsilon \rightarrow 0$ ” if $f(\epsilon) = O(g(\epsilon))$ and $g(\epsilon) = O(f(\epsilon))$ as $\epsilon \rightarrow 0$. The following theorem investigates the expected waiting time until the set of stochastically stable states is first reached. Its proof is provided in the Appendix.⁶

⁶We apply the Radius-(Modified) Coradius theorem from Ellison (2000) to complete the proof of Theorem 2. For $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$, there exist three absorbing sets, and we can only establish an upper bound for the order of the waiting time. To the best of our knowledge, obtaining the exact order necessitates the iterative cycle decomposition (see, e.g., Cui and Zhai (2010) and Newton and Sandholm (2022)). Unfortunately, implementing this technique is cumbersome due to the lengthy mathematical expressions involved.

Theorem 2. Suppose that $k \leq \frac{n-1}{2}$. Then,

$$W(\mathcal{S}^*, \epsilon) = \begin{cases} \Theta(\epsilon^{-c^*}), & \text{if either } n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1 \text{ or } n^h \geq \frac{a-d}{b-d}k; \\ O(\epsilon^{-\min\{c^*, c^{**}\}}), & \text{if } \frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k, \end{cases}$$

as $\epsilon \rightarrow 0$, where

$$c^* = \max \left\{ \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h, \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right\}$$

and

$$c^{**} = \max \left\{ \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, \left\lceil \frac{a-d}{b-d}k \right\rceil - 2 \left(n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right) \right\}.$$

Given that $k \leq \frac{n-1}{2}$, by Theorem 1, it holds that $\mathcal{S}^* = \vec{B}$, which is the same as in Staudigl and Weidenholzer (2014). Note that when all agents are homogeneous with low linking costs (as assumed in Staudigl and Weidenholzer (2014)), the set $\vec{B}$ of stochastically stable states can be reached with a speed of the order of $\left\lceil \frac{a-d}{b-d}k \right\rceil$. By simple calculation, we have

$$c^* = \max \left\{ \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h, \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right\} < \left\lceil \frac{a-d}{b-d}k \right\rceil$$

and

$$c^{**} = \max \left\{ \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, \left\lceil \frac{a-d}{b-d}k \right\rceil - 2 \left(n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right) \right\} < \left\lceil \frac{a-d}{b-d}k \right\rceil$$

provided that k is large. Hence, Theorem 2 establishes that the waiting time to reach the payoff-dominant convention $\vec{B}$ is of a strictly lower order in the presence of heterogeneity among agents, with some agents incurring high linking costs.

The underlying mechanism is as follows. Roughly speaking, agents with high linking costs act as a catalytic minority in the transition from the risk-dominant convention $\vec{A}$ to the payoff-dominant convention $\vec{B}$. First, consider the case where either $n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1$ or $n^h \geq \frac{a-d}{b-d}k$ holds. In this case, there exist two absorbing sets, $\vec{A}$ and $\vec{B}$. Assume that

the dynamics reaches a profile in $\vec{A}$. It is much easier for agents in I^h to switch from the risk-dominant action A to the payoff-dominant action B , in comparison to agents in I^l , since agents in I^h are motivated to form links only with other B -players when they play action B . In fact, when $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ mutants in I^l switch from action A to action B , every agent in I^h will find it optimal to change to action B . Thus, $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ mutations can lead to the emergence of $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil + n^h$ B -players, which may be enough to induce the remaining A -players in I^l to change to action B .⁷ Then, consider the case where $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$ holds. In this case, three absorbing sets, $\vec{A}$, $\vec{AB}$, and $\vec{B}$, exist. In this case, the step-by-step reasoning laid out in Ellison (2000) can apply, with $\vec{AB}$ serving as a stepping stone.

It is natural and insightful to explore the welfare implications of Theorem 2. Consider a homogeneous population in which all agents face the high linking cost γ^h . Then, two absorbing sets arise, $\vec{A}$ and $\vec{B}$, of which only $\vec{B}$ is stochastically stable, with $W(\vec{B}, \epsilon) = \Theta\left(\epsilon^{-\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil}\right)$ as $\epsilon \rightarrow 0$. In other words, the waiting time required to reach the payoff-dominant convention $\vec{B}$ is of order $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$. Note that all agents will secure only a payoff of $k(b - \gamma^h)$ in the long run under the high linking cost γ^h . By contrast, in our setting, when an agent incurs the low linking cost γ^l , she will receive a payoff of $k(b - \gamma^l)$ in the long-run, which is strictly larger than $k(b - \gamma^h)$, and when an agent incurs the high linking cost γ^h , she will receive a payoff of $k(b - \gamma^h)$ in the long-run. Importantly, Theorem 2 demonstrates that a small group of agents with high linking costs can accelerate convergence to the payoff-dominant convention. This observation implies that only limited welfare losses from high linking costs are necessary. More precisely, if n^h increases by one unit, the order of the waiting time decreases by at least one, so long as n^h remains small. In particular, when there exist exactly $\left\lceil \frac{a-d}{b-d}k \right\rceil - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ such agents, $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ mutations may suffice to initiate the transition from $\vec{A}$ to $\vec{B}$.⁸ In addition, an increase in the high linking cost γ^h will create two countervailing effects. On the one hand, both orders identified in Theorem

⁷However, when all agents are homogeneous and have a low linking cost as in Staudigl and Weidenholzer (2014), $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ mutations can only lead to the emergence of $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ B -players, provided that $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil < \left\lceil \frac{a-d}{b-d}k \right\rceil$.

⁸Suppose that $n^h = \left\lceil \frac{a-d}{b-d}k \right\rceil - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$. Straightforward calculations imply that when either $\left\lceil \frac{a-d}{b-d}k \right\rceil \leq \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil + \frac{a-\gamma^h}{b-\gamma^h}k + 1$ or $\left\lceil \frac{a-d}{b-d}k \right\rceil \geq 3 \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$, the bounds derived in Theorem 2 coincide with $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$.

2 will decrease, accelerating convergence toward the payoff-dominant convention. On the other hand, a higher γ^h will inflict larger payoff losses on agents in I^h subject to high linking costs.

Example 1. Consider a coordination game with parameters $[a, b, c, d] = [5, 7, 4, 1]$. Suppose that $n^l = 32$, $n^h = 11$, and $k = 18$. Observe that $k < \frac{n-1}{2} = 21$. Assume $\gamma^h = 3$ and $\gamma^l = 0$. By simple calculation, $\frac{a-\gamma^h}{b-\gamma^h}k = \frac{2}{4} \cdot 18 = 9$ and $\frac{a-d}{b-d}k = \frac{4}{6} \cdot 18 = 12$. The inequality $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$ is satisfied. Accordingly, three absorbing sets emerge: $\vec{A}$, $\vec{AB}$ and $\vec{B}$, among which only $\vec{B}$ is stochastically stable.

By simple calculation,

$$c^* = \max \left\{ \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h, \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right\} = \max\{12 - 11, 9\} = 9$$

and

$$\begin{aligned} c^{**} &= \max \left\{ \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, \left\lceil \frac{a-d}{b-d}k \right\rceil - 2 \left(n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right) \right\} \\ &= \max\{9, 12 - 2(11 - 9)\} = 9. \end{aligned}$$

Then, the modified coradius of the basin of attraction of $\vec{B}$ is given by $\min\{c^*, c^{**}\} = 9$. By Theorem 2, we obtain

$$W(\vec{B}, \epsilon) = O\left(\epsilon^{-\min\{c^*, c^{**}\}}\right) = O(\epsilon^{-9})$$

as $\epsilon \rightarrow 0$. However, when all agents are homogeneous and have a low linking cost, as in Staudigl and Weidenholzer (2014), we have $W(\vec{B}, \epsilon) = \Theta(\epsilon^{-\lceil \frac{a-d}{b-d}k \rceil}) = \Theta(\epsilon^{-12})$ as $\epsilon \rightarrow 0$.

5 Conclusion

We extend the model of Staudigl and Weidenholzer (2014) to explore the role of heterogeneous linking costs. We assume that certain agents incur low costs when forming links, satisfying $0 \leq \gamma^l \leq d$, whereas others face high linking costs such that $d < \gamma^h \leq c$. We

demonstrate the existence of action-heterogeneous Nash equilibria, in which low-linking-cost agents choose the risk-dominant action, whereas high-linking-cost agents choose the payoff-dominant action. This equilibrium arises because high linking costs impose tighter constraints on agents' linking decisions, inducing a stronger incentive to play the payoff-dominant action. Furthermore, under perturbed dynamics, we find that the presence of high-linking-cost agents significantly accelerates convergence to the payoff-dominant convention. Therefore, although heterogeneous linking costs may generate private welfare losses for high-linking-cost agents, they yield positive social benefits and improve coordination efficiency for the entire population.

Appendix A The link-optimized payoff (LOP)

We consider optimal linking decisions while keeping the current choice of action fixed, following the derivation of link-optimized payoffs in Staudigl and Weidenholzer (2014), Cui and Weidenholzer (2021) and Cui and Shi (2022).

For agent i , the link-optimized payoff (LOP) of playing action α_i , denoted by $v_i(\alpha_i, m, \gamma_i)$, is the payoff when agent i with linking cost $\gamma_i \in \{\gamma^l, \gamma^h\}$ forms links optimally given a strategy profile of other agents s_{-i} . Formally,

$$v_i(\alpha_i, m, \gamma_i) = \max_{g_i \in \mathcal{G}_i} \pi_i((\alpha_i, g_i), s_{-i}).$$

Given that $0 \leq \gamma_i \leq c$, every A -agent always receives a nonnegative net payoff regardless of whether she links to an A -agent or a B -agent. Thus, every A -agent is always motivated to form k links. Given $a > c$, each A -agent i prefers to link to other A -agents, but also links to B -agents if she still has other available links. Specifically, the LOP of action A is

$$v_i(A, m, \gamma_i) = a \min\{k, m - 1\} + c(k - \min\{k, m - 1\}) - \gamma_i k.$$

Following the same reasoning, every B -agent $j' \in I^l$ with $0 \leq \gamma_{j'} \leq d$ must form k links. Given $b > d$, agent j' prefers to link to other B -agents, but also links to A -agents if she still has other available links. However, for any B -agent $j'' \in I^h$, given that $d < \gamma_{j''}$, she forms

links only with other B -agents and does not form links with A -agents. Specifically, the LOP of action B is

$$v_j(B, m, \gamma_j) = (b - \gamma_j) \min\{k, n - m - 1\} + \mathbb{1}_{I^I}(j)(d - \gamma_j)(k - \min\{k, n - m - 1\}),$$

where $\mathbb{1}_{I^I}(\cdot)$ is the indicator function.

We finally characterize the conditions under which an agent will choose either of the two actions. An A -agent i strictly prefers action A to action B if $v_i(A, m, \gamma_i) > v_i(B, m - 1, \gamma_i)$, prefers action B to action A if $v_i(A, m, \gamma_i) < v_i(B, m - 1, \gamma_i)$, and is indifferent if $v_i(A, m, \gamma_i) = v_i(B, m - 1, \gamma_i)$. Similarly, a B -agent j prefers action B if $v_j(B, m, \gamma_j) > v_j(A, m + 1, \gamma_j)$, prefers action A if $v_j(B, m, \gamma_j) < v_j(A, m + 1, \gamma_j)$, and is indifferent otherwise.

Appendix B Proofs

B.1 Proof of Lemma 1 and Proposition 1

Proof of Lemma 1. Consider an action-heterogeneous Nash equilibrium s^* . Trivially, the number of B -players is strictly smaller than k . Otherwise, any A -player is strictly motivated to change to action B .

We then show that all the agents in I^h must choose the same action. Assume that there exist two distinct agents i and j in I^h satisfying that agent i plays action A and agent j plays action B . Agent i receives a payoff of $(a - \gamma^h) \min\{k, m - 1\} + (c - \gamma^h)(k - \min\{k, m - 1\})$, which should be larger than $(b - \gamma^h)(n - m)$, the maximal payoff she would obtain when switching to action B , provided that $d < \gamma^h \leq c$. However, agent j receives a payoff of $(b - \gamma^h)(n - m - 1)$, which should be larger than $(a - \gamma^h) \min\{k, m\} + (c - \gamma^h)(k - \min\{k, m\})$, the maximal payoff she would obtain when switching to action A . As a result, we have

$$(a - \gamma^h) \min\{k, m - 1\} + (c - \gamma^h)(k - \min\{k, m - 1\}) \geq (b - \gamma^h)(n - m)$$

and

$$(b - \gamma^h)(n - m - 1) \geq (a - \gamma^h) \min\{k, m\} + (c - \gamma^h)(k - \min\{k, m\})$$

which yield a contradiction. Similarly, all the agents in I^l should choose the same action.

Next, we show that the profile where the agents in I^l choose action B and the agents in I^h choose action A is not a Nash equilibrium. Note that the number of B -players, given by n^l , is strictly smaller than k . We consider two cases: $n^h > k$ and $n^h \leq k$.

First, consider the case in which $n^h > k$. An A -agent $j \in I^h$ can link to k A -agents and receive a payoff of $ak - \gamma^h k$. If agent j switches to action B , she links to all B -agents but not A -agents, because of the high linking cost, $d < \gamma^h \leq c$, and she would obtain a payoff of $bn^l - \gamma^h n^l$. By the definition of Nash equilibrium, we have that

$$ak - \gamma^h k \geq bn^l - \gamma^h n^l.$$

A B -agent $i \in I^l$ can link to all other B -agents and some A -agents. Because of low linking cost, $0 \leq \gamma^l \leq d$, she receives a payoff of $b(n^l - 1) + d(k - n^l + 1) - \gamma^l k$. When agent i switches to action A , she would get a payoff of $ak - \gamma^l k$, by forming k links with agents in I^h . By the definition of Nash equilibrium, we have that

$$b(n^l - 1) + d(k - n^l + 1) - \gamma^l k \geq ak - \gamma^l k.$$

These two inequalities lead to a contradiction.

Then, consider the case in which $n^h \leq k$. An A -agent $j \in I^h$ can link to all other A -agents and some B -agents, she receives a payoff of $a(n^h - 1) + c(k - n^h + 1) - \gamma^h k$. If agent j switches to action B and links to all B -agents, she would receive a payoff of $bn^l - \gamma^h n^l$. By the definition of Nash equilibrium, we have that

$$a(n^h - 1) + c(k - n^h + 1) - \gamma^h k \geq bn^l - \gamma^h n^l.$$

A B -agent $i \in I^l$ can link to all other B -agents and some A -agents, she receives a payoff of

$b(n^l - 1) + d(k - n^l + 1) - \gamma^l k$. When agent i switches to action A , she would obtain a payoff of $an^h + c(k - n^h) - \gamma^l k$ by forming n^h links with all the agents in I^h and $k - n^h$ links with agents in I^l . By the definition of Nash equilibrium, we have that

$$b(n^l - 1) + d(k - n^l + 1) - \gamma^l k \geq an^h + c(k - n^h) - \gamma^l k.$$

These two inequalities yield a contradiction.

Consequently, the strategy profile where the agents in I^l choose action B and the agents in I^h choose action A cannot be a Nash equilibrium. $\square$

Proof of Proposition 1. Clearly, a strategy profile where all agents choose the same action and form all k links is a Nash equilibrium.

Now we turn to the second part. It is sufficient to study the “only if” part. The “if” part can be analyzed similarly. Consider an action-heterogeneous Nash equilibrium s^* . By Lemma 1, all the agents in I^l choose action A , and all the agents in I^h choose action B ; furthermore, $n^h < k$.

First, consider the case where $n^l \leq k$. In this case, each A -agent $i \in I^l$ forms $n^l - 1$ links with A -agents and $k - n^l + 1$ links with B -agents, and receives a payoff of $a(n^l - 1) + c(k - n^l + 1) - \gamma^l k$. If agent i switches to action B , she would receive a payoff of $bn^h + d(k - n^h) - \gamma^l k$, by forming n^h links with all B -agents in I^h and $(k - n^h)$ links with other A -agents in I^l . It follows that

$$a(n^l - 1) + c(k - n^l + 1) - \gamma^l k \geq bn^h + d(k - n^h) - \gamma^l k,$$

which can be rewritten as

$$n^h \leq \frac{(a - c)(n^l - 1) + (c - d)k}{b - d} = \frac{(a - c) \min\{n^l - 1, k\} + (c - d)k}{b - d}. \quad (2)$$

Each B -agent $j \in I^h$ forms links only with the other B -agents, and she receives a payoff of $(b - \gamma^h)(n^h - 1)$, provided that $d < \gamma^h \leq c$. If agent j switches to action A , she would receive a payoff of $an^l + c(k - n^l) - \gamma^h k$, by forming n^l links with A -agents in I^l and $(k - n^l)$ links

with B -agents in I^h . It follows that

$$(b - \gamma^h)(n^h - 1) \geq an^l + c(k - n^l) - \gamma^h k,$$

which can be rewritten as

$$n^h \geq \frac{(a - c)n^l + (c - \gamma^h)k}{b - \gamma^h} + 1 = \frac{(a - c) \min\{n^l, k\} + (c - \gamma^h)k}{b - \gamma^h} + 1. \quad (3)$$

By the inequalities (2) and (3), we obtain

$$\frac{(a - c) \min\{n^l, k\} + (c - \gamma^h)k}{b - \gamma^h} + 1 \leq n^h \leq \frac{(a - c) \min\{n^l - 1, k\} + (c - d)k}{b - d}.$$

Then, consider the case where $n^l \geq k + 1$. In this case, each A -agent $i \in I^l$ forms k links with other A -agents, and receives a payoff of $ak - \gamma^l k$. If agent i switches to action B , she would receive a payoff of $bn^h + d(k - n^h) - \gamma^l k$, by forming n^h links with all B -agents in I^h and $k - n^h$ links with other A -players. It follows that

$$ak - \gamma^l k \geq bn^h + d(k - n^h) - \gamma^l k,$$

which can be written as

$$n^h \leq \frac{a - d}{b - d} k = \frac{(a - c) \min\{n^l - 1, k\} + (c - d)k}{b - d}. \quad (4)$$

In addition, each B -agent $j \in I^h$ only forms links with the other B -agents in I^h , given that $d < \gamma^h \leq c$, and she receives a payoff of $(b - \gamma^h)(n^h - 1)$. If agent j switches to action A , she would receive a payoff of $(a - \gamma^h)k$, by forming k links with A -agents in I^l . It then follows that

$$(b - \gamma^h)(n^h - 1) \geq (a - \gamma^h)k,$$

which can be written as

$$n^h \geq \frac{a - \gamma^h}{b - \gamma^h} k + 1 = \frac{(a - c) \min\{n^l, k\} + (c - \gamma^h)k}{b - \gamma^h} + 1. \quad (5)$$

Inequalities (4) and (5) yield that

$$\frac{(a - c) \min\{n^l, k\} + (c - \gamma^h)k}{b - \gamma^h} + 1 \leq n^h \leq \frac{(a - c) \min\{n^l - 1, k\} + (c - d)k}{b - d}.$$

The proof is completed. $\square$

B.2 Proof of Proposition 2

Proof of Proposition 2. Following reasoning similar to that in the proof of Lemma 1 in Staudigl and Weidenholzer (2014), with positive probability, the unperturbed dynamics can arrive at a state in which all the agents in I^l choose the same action a and all the agents in I^h choose the same action a' . If $a = a'$, it is clear that the dynamics will then move to either $\vec{A}$ or $\vec{B}$.

If $a \neq a'$, by the argument in the proof of Lemma 1, the case in which $a = B$ and $a' = A$ cannot occur. Hence we only need to consider the case in which $a = A$ and $a' = B$. In addition, following reasoning similar to that in the proof of Proposition 1, $\vec{AB}$ is an absorbing set if and only if $\frac{(a-c) \min\{n^l, k\} + (c-\gamma^h)k}{b-\gamma^h} + 1 < n^h < \frac{(a-c) \min\{n^l-1, k\} + (c-d)k}{b-d}$. $\square$

B.3 Proof of Theorems 1 and 2

Consider two absorbing sets Ω and Ω' . The transition cost $c(\Omega, \Omega')$ is defined as the minimal number of mutations required for the transition from Ω to Ω' .

To prove Theorem 1, we introduce two lemmas to determine the minimum number of mutations required for agents in one group to switch to action B , given that the dynamics reaches a state in $\vec{A}$.

Lemma B.1. *Consider the absorbing set $\vec{A}$. The minimum number of mutations required to induce all the agents in I^h to switch to action B is $\min\left\{n^h, \left\lceil \frac{a-\gamma^h}{b-\gamma^h} k \right\rceil\right\}$.*

Proof. Without loss of generality, it is sufficient to explore the case where $n^h > \left\lceil \frac{a-\gamma^h}{b-\gamma^h} k \right\rceil$. The analysis of the other case where $n^h \leq \left\lceil \frac{a-\gamma^h}{b-\gamma^h} k \right\rceil$ is trivial, and is omitted here.

Consider a state $s \in \vec{A}$. We denote by $x^h \leq k$ the minimal number of agents in I switching from action A to action B such that the remaining A -agents in I^h find it optimal to change to action B . Assume that x^h agents get revision opportunities and switch to action B by mistake. Then, for any remaining A -agent $j \in I^h$, if she changes to action B , she will receive a payoff of $(b - \gamma^h)x^h$, which can be attained by forming x^h links with the mutants, given that $\gamma^h > d$. If agent j continues to play action A , she can form links with k A -agents and receive a payoff of $(a - \gamma^h)k$. Agent j is motivated to change to action B whenever $(b - \gamma^h)x^h \geq (a - \gamma^h)k$, or equivalently, $x^h = \left\lceil \frac{a-\gamma^h}{b-\gamma^h} k \right\rceil$. Trivially, fewer mutations cannot induce the A -agents in I^h to change their action. $\square$

The following lemma shows how to induce all the agents in I^l to switch to action B . The proof is similar to the proof of Lemma B.1; therefore, we omit it.

Lemma B.2. *Consider the absorbing set $\vec{A}$. The minimum number of mutations required to induce all the agents in I^l to switch to action B is $\min \left\{ n^l, \left\lceil \frac{a-d}{b-d} k \right\rceil \right\}$.*

Proof of Theorem 1. We use the Radius-(Modified) Coradius theorem in Ellison (2000) to complete the proof. First, we calculate the radius of $\vec{B}$. Note that when $k + 1$ or more B -players exist, it is optimal for every agent to play action B . Therefore, to leave the basin of attraction of $\vec{B}$, at least $n - k$ mutations are needed. By the assumption that $k \leq \frac{n-1}{2}$, we have $n - k \geq \frac{n+1}{2} = \frac{n-1}{2} + 1 \geq k + 1$. That is, $R(\vec{B}) \geq k + 1$.

We now proceed to compute the modified coradius of $\vec{B}$. Given that $n^h < k$, the inequality $n^l > k$ always holds, provided that $k \leq \frac{n-1}{2}$. Notably, $\frac{(a-c) \min\{n^l, k\} + (c-\gamma^h)k}{b-\gamma^h} + 1 < n^h < \frac{(a-c) \min\{n^l-1, k\} + (c-d)k}{b-d}$ is equivalent to $\frac{a-\gamma^h}{b-\gamma^h} k + 1 < n^h < \frac{a-d}{b-d} k$. In accordance with the characterization of absorbing sets in Proposition 2, the analysis is divided into two cases, namely, the case where either $n^h \leq \frac{a-\gamma^h}{b-\gamma^h} k + 1$ or $n^h \geq \frac{a-d}{b-d} k$ and the case where $\frac{a-\gamma^h}{b-\gamma^h} k + 1 < n^h < \frac{a-d}{b-d} k$, by introducing the following two lemmas.

Lemma B.3. Suppose either $n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1$ or $n^h \geq \frac{a-d}{b-d}k$. Then,

$$CR(\vec{B}) = \begin{cases} \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, & \text{if } n^h + \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \geq \left\lceil \frac{a-d}{b-d}k \right\rceil; \\ \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h, & \text{if } n^h + \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil < \left\lceil \frac{a-d}{b-d}k \right\rceil. \end{cases}$$

Moreover, $CR(\vec{B}) = \max \left\{ \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h \right\} = c^*$, and $CR(\vec{B}) \leq \left\lceil \frac{a-d}{b-d}k \right\rceil$ always holds.

Proof. By Proposition 2, there are two absorbing sets: $\vec{A}$ and $\vec{B}$, provided that either $n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1$ or $n^h \geq \frac{a-d}{b-d}k$. It is sufficient to consider the case where $n^h + \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil < \left\lceil \frac{a-d}{b-d}k \right\rceil$. The other case can be analyzed similarly.

Consider a state $s \in \vec{A}$. Assume that $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ agents in the set I^l make mistakes and change from action A to action B . Then, by Lemma B.1, every agent in I^h is motivated to choose action B rather than action A and form links with the mutants. Suppose that only these n^h agents in I^h receive revision opportunities. As a result, we can reach the state where both the mutants and all the agents in I^h choose action B . Then, the number of B -players is $n^h + \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$.

Note that $n^h + \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil < \left\lceil \frac{a-d}{b-d}k \right\rceil$. By Lemma B.2, to ensure that every remaining A -player, who must belong to the set I^l , is motivated to change her action choice, more mutations are needed. Assume that $\left\lceil \frac{a-d}{b-d}k \right\rceil - n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ A -players make mistakes and switch from action A to action B . This forces the number of B -players to be $\left\lceil \frac{a-d}{b-d}k \right\rceil$. Applying Lemma B.2 again shows that every remaining A -agent in I^l now finds it optimal to change her action. Then, the absorbing set $\vec{B}$ can be obtained.

Consequently, we have

$$\begin{aligned} CR(\vec{B}) &= \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil + \left(\left\lceil \frac{a-d}{b-d}k \right\rceil - n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right) \\ &= \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h. \end{aligned}$$

It is straightforward that to transit from $\vec{A}$ to $\vec{B}$, at least $\left\lceil \frac{a-d}{b-d}k \right\rceil - n^h$ mutations are required. Trivially, $CR(\vec{B}) < \left\lceil \frac{a-d}{b-d}k \right\rceil$. Furthermore, via simple calculation, $CR(\vec{B})$ can be rewritten as $\max \left\{ \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, \left\lceil \frac{a-d}{b-d}k \right\rceil - n^h \right\}$. $\square$

Lemma B.4. Suppose that $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$. Then, $CR^*(\vec{B}) = \min\{c^*, c^{**}\}$, where

$$c^{**} = \max \left\{ \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil, \left\lceil \frac{a-d}{b-d}k \right\rceil - 2 \left(n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil \right) \right\}.$$

Proof. By Proposition 2, there are three absorbing sets: $\vec{A}$, $\vec{B}$, and $\vec{AB}$, provided that $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$.

To complete the proof, we consider two cases. In the first case, the least costly path from $\vec{A}$ to $\vec{B}$ leaves the basin of attraction of $\vec{A}$ and then enters the basin of attraction of $\vec{B}$ directly, without passing through $\vec{AB}$. Following the same reasoning as in the proof of Lemma B.3, in this case, the modified cost of the least costly path is c^* .

In the second case, the least costly path from $\vec{A}$ to $\vec{B}$ passes through $\vec{AB}$. In this case, the analysis is divided into three steps.

In the first step, consider the transition from $\vec{A}$ to $\vec{AB}$. By the inequality $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h$, it holds that $n^h \geq \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil + 1$. Consider a state $s \in \vec{A}$. Assume that $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ agents in I^h make mistakes and switch from action A to action B . By Lemma B.1, every remaining A -player in I^h is motivated to change to action B . Assume that only these $n^h - \left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ remaining A -players in I^h receive revision opportunities. Then, the absorbing set $\vec{AB}$ can be obtained. Thus, $\left\lceil \frac{a-\gamma^h}{b-\gamma^h}k \right\rceil$ mutations can induce the transition from $\vec{A}$ to $\vec{AB}$.

In the second step, consider the transition from $\vec{AB}$ to $\vec{B}$. Consider a state $s' \in \vec{AB}$. Note that for any state in $\vec{AB}$, the number of B -players is given by n^h . By the inequality $n^h < \frac{a-d}{b-d}k$, it holds that $n^h < \left\lceil \frac{a-d}{b-d}k \right\rceil$. Assume that $\left\lceil \frac{a-d}{b-d}k \right\rceil - n^h$ agents in I^l make mistakes and switch from action A to action B . As a result, the number of B -players turns out to be $\left\lceil \frac{a-d}{b-d}k \right\rceil$. By Lemma B.2, every remaining A -player in I^l has an incentive to change to action B . Assume that only these $n - \left\lceil \frac{a-d}{b-d}k \right\rceil$ remaining A -players in I^l get revision opportunities. Then, the absorbing set $\vec{B}$ can be reached. Hence, $\left\lceil \frac{a-d}{b-d}k \right\rceil - n^h$ mutations can lead to the transition from $\vec{AB}$ to $\vec{B}$.

Note that at each of the above steps, we identify the least costly way to accomplish the

transition. Consequently, we have

$$\begin{aligned} c(\vec{A}, \vec{B}) &= c(\vec{A}, \vec{AB}) + c(\vec{AB}, \vec{B}) \\ &= \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil + \left(\left\lceil \frac{a - d}{b - d} k \right\rceil - n^h \right). \end{aligned}$$

Finally, we consider the modified cost of the transition from $\vec{A}$ to $\vec{B}$, denoted by $c^*(\vec{A}, \vec{B})$. Note that $c^*(\vec{A}, \vec{B}) = c(\vec{A}, \vec{B}) - R(\vec{AB})$, where $R(\vec{AB})$ is the radius of the basin of attraction of the absorbing set $\vec{AB}$. We now prove that

$$R(\vec{AB}) = \min \left\{ \left\lceil \frac{a - d}{b - d} k \right\rceil - n^h, n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil \right\}.$$

First, we examine the transition from $\vec{AB}$ to $\vec{A}$. Consider a state s'' in $\vec{AB}$, in which the number of B -players is n^h . Assume that $n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil$ agents in I^h make mistakes and change from action B to action A . Then, the number of B -players decreases to $\left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil$. Following the same reasoning as in the proof of Lemma B.1, every remaining B -player in I^h now finds it optimal to switch to action A since only $\left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil - 1$ other agents play action B . In addition, as a direct consequence of Lemma B.1, $n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil - 1$ or fewer mutations fail to induce any B -player to change action. Thus, at least $n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil$ mutations are required to induce the transition from $\vec{AB}$ to $\vec{A}$. Moreover, it follows from the foregoing analysis of the transition from $\vec{AB}$ to $\vec{B}$ that at least $\left\lceil \frac{a - d}{b - d} k \right\rceil - n^h$ mutations must occur to accomplish this transition. Hence, $R(\vec{AB}) = \min \left\{ \left\lceil \frac{a - d}{b - d} k \right\rceil - n^h, n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil \right\}$.

Thus, we have

$$\begin{aligned} c^*(\vec{A}, \vec{B}) &= c(\vec{A}, \vec{B}) - R(\vec{AB}) \\ &= \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil + \left(\left\lceil \frac{a - d}{b - d} k \right\rceil - n^h \right) - \min \left\{ \left\lceil \frac{a - d}{b - d} k \right\rceil - n^h, n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil \right\} \\ &= \max \left\{ \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil, \left\lceil \frac{a - d}{b - d} k \right\rceil - 2 \left(n^h - \left\lceil \frac{a - \gamma^h}{b - \gamma^h} k \right\rceil \right) \right\}. \end{aligned}$$

Note that the modified coradius of the basin of attraction of $\vec{B}$ is given by the minimum of the costs identified in each case. Formally, $CR^*(\vec{B}) = \min\{c^*, c^{**}\}$. $\square$

Note that when there are exactly two absorbing sets, the coradius coincides with its modified counterpart. By Lemmas B.3 and B.4, in either case, the minimum resistance of the transition from any other absorbing set to $\vec{B}$ is no more than $\lceil \frac{a-d}{b-d}k \rceil$; i.e., $CR^*(\vec{B}) \leq \lceil \frac{a-d}{b-d}k \rceil$. Trivially, $\lceil \frac{a-d}{b-d}k \rceil \leq k$ holds. Clearly, $R(\vec{B}) > CR^*(\vec{B})$, given that $R(\vec{B}) \geq k+1$. By Theorem 2 in Ellison (2000), $\mathcal{S}^* = \vec{B}$. $\square$

Proof of Theorem 2. By Theorem 1, when $k \leq \frac{n-1}{2}$, the set of stochastically stable states is $\vec{B}$, i.e., $\mathcal{S}^* = \vec{B}$. To study the waiting time, depending on the value of n^h , we distinguish three cases: (i) $n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1$, (ii) $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$, and (iii) $n^h \geq \frac{a-d}{b-d}k$.

First, consider case (i) $n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1$. By Proposition 2, in this case, there are two absorbing sets: $\vec{A}$ and $\vec{B}$. From the proof of Theorem 1, the radius of the basin of attraction of $\vec{B}$ satisfies $R(\vec{B}) \geq k+1$. In addition, Lemma B.3 shows that the coradius of the basin of attraction of $\vec{B}$ is given by $\max\left\{\lceil \frac{a-d}{b-d}k \rceil - n^h, \lceil \frac{a-\gamma^h}{b-\gamma^h}k \rceil\right\} \leq \lceil \frac{a-d}{b-d}k \rceil$. It holds that $CR(\vec{B}) \leq k$. As an application of Theorem 1 in Ellison (2000), when $n^h \leq \frac{a-\gamma^h}{b-\gamma^h}k + 1$, $W(s, \vec{B}, \epsilon) = O\left(\epsilon^{-\max\left\{\lceil \frac{a-d}{b-d}k \rceil - n^h, \lceil \frac{a-\gamma^h}{b-\gamma^h}k \rceil\right\}}\right)$ as $\epsilon \rightarrow 0$ for all $s \notin D(\vec{B})$, where $D(\vec{B})$ is the basin of attraction of $\vec{B}$. Furthermore, by the proof of Lemma B.3, to leave the basin of attraction of $\vec{A}$, at least $\max\left\{\lceil \frac{a-d}{b-d}k \rceil - n^h, \lceil \frac{a-\gamma^h}{b-\gamma^h}k \rceil\right\}$ mutations are required. Hence, it is true that $W(\vec{B}, \epsilon) = \max_{s \in \mathcal{S} \setminus D(\vec{B})} W(s, \vec{B}, \epsilon) = \Theta\left(\epsilon^{-\max\left\{\lceil \frac{a-d}{b-d}k \rceil - n^h, \lceil \frac{a-\gamma^h}{b-\gamma^h}k \rceil\right\}}\right) = \Theta(\epsilon^{-c^*})$ as $\epsilon \rightarrow 0$.

Next, we study case (ii) $\frac{a-\gamma^h}{b-\gamma^h}k + 1 < n^h < \frac{a-d}{b-d}k$. In this case, Proposition 2 establishes that there are three absorbing sets: $\vec{A}$, $\vec{AB}$, and $\vec{B}$. From the proof of Theorem 1, the radius of the basin of attraction of $\vec{B}$ satisfies $R(\vec{B}) \geq k+1$. Additionally, Lemma B.4 indicates that the modified coradius of the basin of attraction of $\vec{B}$ is $CR^*(\vec{B}) = \min\{c^*, c^{**}\} < k$. By applying Theorem 2 in Ellison (2000), we have $W(s, \vec{B}, \epsilon) = O\left(\epsilon^{-\min\{c^*, c^{**}\}}\right)$ as $\epsilon \rightarrow 0$ for all $s \notin D(\vec{B})$.

Finally, we turn to case (iii) $n^h \geq \frac{a-d}{b-d}k$. This case can be analyzed using the same reasoning as case (i). Notably, when $n^h \geq \frac{a-d}{b-d}k$, it must be true that $\lceil \frac{a-\gamma^h}{b-\gamma^h}k \rceil \geq \lceil \frac{a-d}{b-d}k \rceil - n^h$. This implies that in this case, the coradius of the basin of attraction of $\vec{B}$ is $\lceil \frac{a-\gamma^h}{b-\gamma^h}k \rceil$. $\square$

References

- Alós-Ferrer, C. and F. Shi (2012). “Imitation with asymmetric memory”. In: *Economic Theory* 49, pp. 193–215.
- Alós-Ferrer, C. and S. Weidenholzer (2008). “Contagion and efficiency”. In: *Journal of Economic Theory* 143, pp. 251–274.
- (2014). “Imitation and the role of information in overcoming coordination failures”. In: *Games and Economic Behavior* 87, pp. 397–411.
- Bala, V. and S. Goyal (2000). “A non-cooperative model of network formation”. In: *Econometrica* 68, pp. 1181–1230.
- Bilancini, E. and L. Boncinelli (2018). “Social coordination with locally observable types”. In: *Economic Theory* 65.4, pp. 975–1009.
- Blume, L. E. (1993). “The statistical mechanics of strategic interaction”. In: *Games and Economic Behavior* 5, pp. 387–424.
- Cui, Z. (2014). “More neighbors, more efficiency”. In: *Journal of Economic Dynamics & Control* 40, pp. 103–115.
- (2023). “Linking friction, social coordination and the speed of evolution”. In: *Games and Economic Behavior* 140, pp. 410–430.
- Cui, Z. and F. Shi (2022). “Bandwagon effects and constrained network formation”. In: *Games and Economic Behavior* 134, pp. 37–51.
- Cui, Z. and R. Wang (2016). “Collaboration in networks with randomly chosen agents”. In: *Journal of Economic Behavior & Organization* 129, pp. 129–141.
- Cui, Z. and S. Weidenholzer (2021). “Lock-in through passive connections”. In: *Journal of Economic Theory* 192. Article 105187, p. 105187.
- Cui, Z. and J. Zhai (2010). “Escape dynamics and equilibria selection by iterative cycle decomposition”. In: *Journal of Mathematical Economics* 46, pp. 1015–1029.
- Ellison, G. (1993). “Learning, local interaction, and coordination”. In: *Econometrica* 61, pp. 1047–1071.
- (2000). “Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution”. In: *Review of Economic Studies* 67, pp. 17–45.

- Foster, D. and H. Young (1990). “Stochastic Evolutionary Game Dynamics”. In: *Theoretical Population Biology* 38, pp. 219–232.
- Freidlin, Mark I. and Alexander D. Wentzell (1984). *Random Perturbations of Dynamical Systems*. Vol. 260. Grundlehren der mathematischen Wissenschaften. New York: Springer-Verlag.
- Goyal, S. and F. Vega-Redondo (2005). “Network formation and social coordination”. In: *Games and Economic Behavior* 50, pp. 178–207.
- Harsanyi, John C. and Reinhard Selten (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass.: MIT Press.
- Jackson, M. O. and A. Watts (2002). “The evolution of social and economic networks”. In: *Journal of Economic Theory* 106. Part a, pp. 265–295.
- Kandori, M., G. Mailath, and R. Rob (1993). “Learning, mutation, and long run equilibria in games”. In: *Econometrica* 61, pp. 29–56.
- Lu, F. and F. Shi (2023). “Coordination with heterogeneous interaction constraints”. In: *Games and Economic Behavior* 142, pp. 645–665.
- Morris, S. (2000). “Contagion”. In: *Review of Economic Studies* 67, pp. 57–78.
- Naono, M. (2022). “Cost heterogeneity and the persistence of bilingualism”. In: *Games and Economic Behavior* 136, pp. 325–339.
- Neary, P. R. (2012). “Competing conventions”. In: *Games and Economic Behavior* 76, pp. 301–328.
- Newton, J. (2021). “Conventions under Heterogeneous Behavioural Rules”. In: *The Review of Economic Studies* 88.4, pp. 2094–2118.
- Newton, J. and S. D. Angus (2015). “Coalitions, tipping points and the speed of evolution”. In: *Journal of Economic Theory* 157, pp. 172–187.
- Newton, J. and W. Sandholm (2022). “Stochastic dynamics and Edmonds’ algorithm”. In: *Journal of Dynamics and Games* 9.4, pp. 529–546.
- Norman, T. W. L. (2009). “Rapid evolution under inertia”. In: *Games and Economic Behavior* 66.2, pp. 865–879.
- Peski, Marcin (2010). “Generalized risk-dominance and asymmetric dynamics”. In: *Journal of Economic Theory* 145.1, pp. 216–248.

- Staudigl, M. and S. Weidenholzer (2014). “Constrained interactions and social coordination”.
In: *Journal of Economic Theory* 152, pp. 41–63.
- Young, H. (1993). “The evolution of conventions”. In: *Econometrica* 61, pp. 57–84.
- (1998). *Individual Strategy and Social Structure*. Princeton, New Jersey: Princeton University Press.