Games and Economic Behavior 155 (2026) 228–249

Contents lists available at ScienceDirect

Games and Economic Behavior

journal homepage: www.elsevier.com/locate/geb

Indignation and the evolution of cooperation norms

Xueheng Li [1]

Lingnan College, Sun Yat-Sen University, Guangzhou, China
Laboratory of Mezzoeconomics and Regional Industrial Coordinated Development, Sun Yat-Sen University, Guangzhou, China

A R T I C L E I N F O

JEL classification:
C73
D63
H41

Keywords:
Social norm
Emotions
Cooperation
Punishment
Stochastic stability
Psychological game

A B S T R A C T

Sociologists and psychologists have long argued that emotions are essential to sustaining social norms. This study examines the role of indignation in upholding cooperation norms within society. I model indignation in a population psychological game and characterize the stochastically stable equilibrium in a noisy best-reply dynamic. The analysis yields two findings. First, indignation can sustain cooperation in the long run, irrespective of whether interactions are global or occur within a fixed local interaction structure. Second, mobility allows individuals to sort into cooperative communities, causing these communities to grow in size and persist under a wider range of conditions. Therefore, mobility fosters stable norms of cooperation and punishment in large human societies. This study is the first to apply stochastic stability to address multiple equilibria in psychological games.

1. Introduction

> "Thus if I believe that another has violated my interest, I may feel anger; if I believe that in doing so he has violated a norm, I feel indignation."
>
> - Jon Elster, *Emotions and Economic Theory* (1998)

The BBC had a TV series called Very British Problems.[1] One such problem is the British obsession with queuing. As George Mikes describes in his book *How to be an Alien* (1946), "An Englishman, even if he is alone, forms an orderly queue of one." Queuing presents a social dilemma for individuals: people can gain from jumping a queue, but doing so would harm all those behind them. Thus, maintaining a queue is a cooperative action: it is costly to individuals but preserves social order and benefits the community as a whole.

Cooperation is a fundamental aspect of human societies. Like queuing in Britain, cooperation is sometimes not sustained by the shadow of the future or enforced by a central authority. Instead, it is sustained as a social norm, one that "…must be shared by other

E-mail address: lixueheng@mail.sysu.edu.cn

[1] I thank Silvia Sonderegger, Alex Possajennikov, and Yann Bramoullé for their advices that have led to significant improvements of this paper, and thank Robert Sugden for insightful discussions on related topics. I also thank Zhiwei Cui and the anonymous referees for many constructive comments. I acknowledge the National Natural Science Foundation of China (grant number 72203105) and the Economic and Social Research Council (grant number ES/K002201/1) via the Network for Integrated Behavioural Science program for financial support. Previous versions of this paper were circulated under the title "Resentment and the Evolution of Cooperative Norms."

[1] Also see the X (Twitter) account @SoVeryBritish, and the book Very British Problems: Making Life Awkward for Ourselves, One Rainy Day at a Time by Rob Temple.

https://doi.org/10.1016/j.geb.2025.10.009
Received 7 April 2024
Available online 2 November 2025

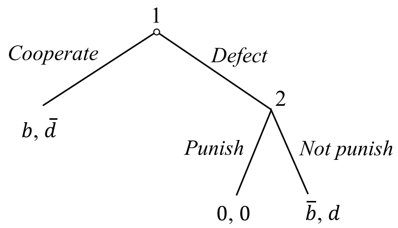

Fig. 1. Individuals in the population are randomly paired to play this cooperation-punishment base game. Numbers at terminal nodes are material payoffs, with $\bar{b} > b > 0$ and $\bar{d} > d > 0$.

people and partly sustained by their approval and disapproval …by the feelings of embarrassment, anxiety, guilt and shame that a person suffers at the prospect of violating them." (Elster, 1989) Thus, it is important to provide a microfoundation for social norms in terms of individual emotions.

In this paper, I propose a model to study the evolution of social norms governing cooperation. These norms, and the willingness to uphold them, are endogenous. Formalizing Elster (1998)'s statement about indignation against norm violations, I assume individuals experience negative feelings and are impelled to punish another person if, and only if, two conditions are met: (*i*) the person's behavior violates their interest, and (*ii*) it also breaches expected common social practices. My focus is on the problem of cooperation and punishment in a social dilemma situation. In both the static game and the noisy best-reply dynamics, indignation creates a complementarity between cooperation and the punishment of defectors. The finding is that this complementarity can be so strong that it not only sustains cooperation in the short run but also leads to the emergence and persistence of cooperation in the long run. This prediction contrasts with a class of previous conformity-based cultural evolutionary models of cooperation (Boyd and Richerson, 1985; Henrich and Boyd, 1998, 2001), where defection would dominate in the long run under the same dynamic framework.

The model also yields different predictions regarding the effects of interaction structures and mobility on cooperation compared to a typical imitation-based evolutionary model of cooperation. An imitation-based model would predict that cooperation is most likely to emerge and spread when interactions occur within local and fixed neighborhoods, and is difficult to emerge under global random matching or in societies with high mobility (Eshel et al., 1998; Nowak and May, 1992; Nowak et al., 2010; Abramson and Kuperman, 2001; Santos and Pacheco, 2005; Szabó and Fáth, 2007; Ohtsuki et al., 2006). My indignation-norm-based account of cooperation implies the opposite: the parameter range to support the long-run stability of the cooperation equilibrium under local interactions is (asymptotically) the same as that under global interactions. This suggests that cooperation is equally likely to emerge and persist under these two interaction structures, although the speed of convergence may differ. Moreover, allowing for free movement between communities promotes the emergence, expansion, and persistence of cooperative communities, rather than inhibiting them.

Formally, indignation is a belief-dependent preference. Whether individual A has indignation against individual B depends on both B's action and A's beliefs regarding the distribution of strategies of all individuals in society. The formalization adapts Battigalli et al. (2019)'s model of anger and frustration to a population game played by all members of society. To examine cooperation and punishment, I focus on the cooperation-punishment base game in Fig. 1.[2,3] Indignation predicts that individuals are willing to incur personal costs to punish defectors if and only if they believe that a sufficient majority cooperates-that cooperation is "the norm." This prediction is supported by experimental evidence.[4] As is common in psychological games (e.g., Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Battigalli et al., 2019), the belief-dependence leads to multiple equilibria (Proposition 1). One equilibrium supports cooperation: if everyone expects that everyone cooperates, then all individuals would punish defectors, and they would cooperate to avoid being punished, which in turn confirms everyone's expectation of cooperation. This equilibrium marks the establishment of a *cooperation norm*. The other is the defection equilibrium: if no one expects cooperation, then there is no indignation and no punishment for defectors, so everyone defects.

The question is whether, and when, the cooperation norm is more likely to emerge and persist than the defection norm. While prior studies on psychological games are restricted to static equilibrium analysis, I extend the stochastic dynamic analysis of Foster and Young (1990), Young (1993), Kandori et al. (1993), Ellison (1993, 2000) to evaluate the long-run stability of each psychological equilibrium. I consider a noisy best-reply dynamic, in which individuals (i) experience indignation and take best responses to their beliefs in each period, (ii) form beliefs by sampling the strategy profile from the previous period, and (iii) make mistakes (i.e., take non-best responses) with a small but positive probability. A state of the dynamic specifies each individual's strategy and their beliefs, corresponding to a strategy-belief profile of the static game. A state is *long-run stable*, or *stochastically stable*, if it occurs with a positive

[2] For most of our analysis, the cooperation-punishment game has a similar material payoff structure as a mini-ultimatum game as studied theoretically by, e.g., Huck and Oechssler (1999) and Battigalli et al. (2019). There are numerous experimental studies on the ultimatum game and its variants; see, e.g., Güth et al. (1982), Camerer and Thaler (1995), Henrich (2000), Engelmann and Strobel (2004), Cooper and Dutcher (2011). In Section 5.2, I impose an additional assumption to derive Theorem 3, namely, that cooperation by player 1 is the unique efficient outcome: $b + \bar{d} > \bar{b} + d$. This distinguishes it from an ultimatum game.

[3] A general model of indignation is presented in the Online Appendix.

[4] Experimental studies show that the predicted behavior is prevalent among participants (Li et al., 2021; Carpenter and Matthews, 2012). For example, in an experiment involving nearly a thousand adult participants, Li et al. (2021) find that a significant fraction would increase their punishment of defectors in the presence of a higher proportion of cooperators.

frequency in the stationary distribution as the probability of mistakes vanishes (Young, 1993; Kandori et al., 1993; Foster and Young, 1990). Based on Ellison (2000), a state will be long-run stable if it is relatively difficult to move away from it and simultaneously relatively easy to reach from all other states, where the difficulty is measured by the minimum number of mistakes necessary for relevant transitions.

Two results emerge from the analysis. First, there are two thresholds. Individuals will cooperate if and only if they believe that the proportion of punishers in their neighborhood is equal to or greater than a threshold γ (between zero and one). They will punish defectors if they believe that the proportions of cooperators and punishers are both equal to or greater than a threshold η (between zero and one). The cooperation equilibrium will be uniquely long-run stable if and only if the sum of η and (a variant of) γ is strictly less than one, irrespective of whether interactions are global or occur in a fixed circle network (Theorems 1 and 2).[5] This suggests that, at least for sufficiently large neighborhoods under symmetric local interaction structures, local versus global interaction structures do not affect the selection of the cooperation norm in the long run.

Second, mobility-the free movement between communities-can promote cooperation through two channels. The first channel is an intermediate *voting-with-feet* effect: once the cooperation norm is established in a community, all individuals free to move will gravitate towards the cooperative community, causing it to grow. The second channel is the emergence of a new set of mixed states where individuals cooperate in large communities. In a long-run stable mixed state, the cooperation norm is established in a dominant large community, which attracts all individuals free to move, while the defection norm prevails in small communities. Such mixed states can be long-run stable if and only if there is free movement between communities; otherwise, the defection equilibrium is long-run stable under the same set of parameters (Theorem 3).

1.1. Related literature

This paper contributes to the study of social norms and cooperation in two ways. First, it formalizes indignation in a population psychological game. Second, it combines the model of indignation with stochastic dynamics and stability analysis to provide a psychologically grounded social evolution model of cooperation.

This study builds on a significant literature on how descriptive norms shape normative expectations and sanctioning behavior (e.g., Sugden, 2000, 1986; Bicchieri, 2006; Carpenter and Matthews, 2012). Building on Humean ideas, Elster (1998) conceptualizes indignation as a reactive emotion triggered by norm violations. Sugden (2000) formalizes this in a prisoner's dilemma context, modeling "resentment" comparable to guilt aversion (Battigalli and Dufwenberg, 2007).

This paper extends this tradition by embedding indignation in a population game where reference payoffs depend on endogenous beliefs about societal norms. This is achieved by adapting Battigalli et al. (2019)'s model of anger and frustration to a population game setup. A crucial departure from Battigalli et al. is that the triggers for indignation are not beliefs concerning the behavior of immediate interaction partners but the broader population. This departure is not only conceptual but also has behavioral consequences as discussed in the remarks in Section 2.1.

My model of indignation also differs markedly from existing models of reciprocity (Gintis, 2000; Fehr et al., 2002; Rabin, 1993; Dufwenberg and Kirchsteiger, 2004), which assume context-*independent* reference points for punishment. Under reciprocity, defection is considered unkind regardless of social context, predicting invariant punishment across different societal levels of cooperation. In contrast, my model predicts that punishment intensifies with higher cooperation levels in society.

While G. Akerlof (1982)'s influential gift-exchange theory recognizes the importance of market-wide reference points, it focuses solely on firm-worker relationships rather than general norm evolution. Similarly, although R. Akerlof (2016)'s model of anger-driven enforcement shares a similar intuition with indignation in that punishment depends on endogenous beliefs about reasonable behavior, it assumes the existence of an exogenous reasonable person type. Consequently, when compliance costs are low, it rules out noncompliance as an equilibrium. In my model, the defection state is always an equilibrium; multiplicity of equilibria is generally an issue.

More importantly, without a dynamic analysis, R. Akerlof (2016)'s model, models of reciprocity, and other existing psychological game studies provide limited insight into equilibrium selection when multiple equilibria arise. In contrast, the current study combines psychological game theory and stochastic dynamics to present a theory that examines the roots of the spontaneous emergence and selection of cooperation norms.[6]

As mentioned at the beginning and detailed in later analyses, my model presents a complementary approach and generates different predictions to existing imitation-based evolutionary models of cooperation (Nowak and May, 1992; Eshel et al., 1998; Abramson and Kuperman, 2001; Santos and Pacheco, 2005; Szabó and Fáth, 2007; Nowak et al., 2010; Ohtsuki et al., 2006) and conformity models for cultural transmissions (Boyd and Richerson, 1985; Henrich and Boyd, 1998, 2001).

Finally, this study complements models on intergenerational cultural transmissions (Bisin and Verdier, 2000, 2001, 2011; Adriani and Sonderegger, 2009; Tabellini, 2008). Bisin and Verdier (2001)'s seminal work considers cultural traits as purposeful socialization decision outcomes of parents and focuses on identifying conditions for cultural heterogeneity within a society. This contrasts with the current study's focus on horizontal cultural transmissions through the complementarities between actions, beliefs, and emotions.

[5] The exact conditions for long-run stability of the cooperation equilibrium under global and local interactions are asymptotically equivalent as neighborhood size goes to infinity.

[6] Norman (2009) and Khan (2022) have also examined stochastic stability in games with nonstandard preferences.

Section 2 presents the model of indignation and the cooperation-punishment game. Section 3 introduces the stochastic dynamics and defines a long-run stable equilibrium. Section 4 characterizes the long-run stable equilibrium in global interactions. Section 5 examines local interactions and mobility. Section 6 concludes.

2. Static model

2.1. The model

I view social norms as commonly expected and widely followed behaviors within society, sometimes referred to as descriptive social norms in the literature (Bicchieri, 2006). The belief-dependent emotion of indignation upholds the norms of cooperation. The model of indignation adapts Battigalli et al. (2019)'s model of anger and frustration to a population game setup. The difference from Battigalli et al. (2019), and the difference between anger and indignation, is that, as Elster (1998) contrasts, indignation is not a feeling of being harmed but a feeling of *being wronged* because the other person takes action that violates both my interests *and* my perception of the social norms. The population game captures this notion.

Consider a population $N = \{1, 2, \dots, n\}$ with a large number of individuals who are randomly matched to play a *base game*. The general model for an arbitrary extensive-form base game of perfect information is presented in the Online Appendix. For the purpose of this study, it suffices to focus on the cooperation-punishment base game displayed in Fig. 1. Each player in the base game is called a *player role*. There are two player roles in the cooperation-punishment game: Player 1 chooses whether to cooperate or defect. If player 1 defects, then player 2 chooses whether to punish player 1. The numbers (π_1, π_2) at the terminal nodes represent material payoffs. Assume $\bar{b} > b > 0$ and $\bar{d} > d > 0$. Thus, punishment is costly.[7]

Player roles are randomly and uniformly assigned. Thus, a *strategy* for each $i \in N$ is a pair $s_i = (x_i, y_i)$, where $x_i = 1$ stands for cooperate, $x_i = 0$ for defect, $y_i = 1$ for punish, and $y_i = 0$ for not punish. Let $x = (x_1, \dots, x_n)$ and $y = (y_1, \dots, y_n)$. Let $s = (s_1, \dots, s_n)$, or $s = (x, y)$, denote the strategy profile of all individuals, and let S be the set of all strategy profiles.

Individual i's *belief* is denoted by a pair (q_i, p_i), where $q_i \in [0, 1]$ is i's belief about the proportion of individuals who choose to cooperate in the role of player 1 and $p_i \in [0, 1]$ is i's belief about the proportion of individuals who choose to punish in the role of player 2 in the population. Let $q = (q_1, \dots, q_n)$ and $p = (p_1, \dots, p_n)$. The (population) *state* is $z = (x, y, q, p)$, which will be the state variable of the dynamics considered later. Let Z be the set of all possible states.

Decision utilities that incorporate indignation are defined as follows. Let $H \equiv \{\varnothing, defect\}$ include the decision nodes of the base game, where $\varnothing$ denotes the start of the base game and *defect* denotes an individual in player role 2 encountering a defector. Suppose individual $i \in N$ is assigned player role 1 and decides whether to cooperate at $\varnothing$. Let $\mathbb{E}[\pi_1|(\varnothing, x_i), q_i, p_i]$ denote individual i's expectation about her material payoffs when choosing $x_i \in \{0, 1\}$. Let $u_i(x_i|\varnothing, q_i, p_i)$ denote i's decision utility at $\varnothing$ when choosing $x_i \in \{0, 1\}$, conditional on her belief. As is clear in the general model in the Online Appendix, a general property of indignation, as well as anger in Battigalli et al. (2019), is that it is an ex-post phenomenon. It never occurs at the start of a game regardless of beliefs. Hence, individuals in player role 1 aim to maximize expected material payoffs:

$$u_i(x_i|\varnothing, q_i, p_i) = \mathbb{E}[\pi_1|(\varnothing, x_i), q_i, p_i].$$

Individuals in player role 2, however, may experience indignation when encountering defectors. Suppose individuals $i \in N$ and $j \in N$ are matched, where i is a player 2 and j is a player 1. Let $\mathbb{E}[\pi_2|\varnothing, q_i, p_i]$ denote i's expected material payoffs at the start of the game if both individuals act according to i's belief about the population norm, i.e., j cooperates with probability q_i and i punishes with probability p_i. That is, $\mathbb{E}[\pi_2|\varnothing, q_i, p_i] = q_i\bar{d} + (1 - q_i)(1 - p_i)d$. Individual i considers these the reference payoffs she is entitled to receive. However, suppose j defects. Conditional on j's defection, individual i's expected payoffs become $\mathbb{E}[\pi_2|defect, q_i, p_i] = (1 - p_i)d$ if she continues to act according to her perception of the norm. I define individual i's *indignation* in player role 2 when she encounters a defector j conditional on her belief (q_i, p_i) by

$$\rho_{21}(q_i, p_i) = \mathbb{E}[\pi_2|\varnothing, q_i, p_i] - \mathbb{E}[\pi_2|defect, q_i, p_i].$$

Given $\bar{d} > d$, we have $\mathbb{E}[\pi_2|\varnothing, q_i, p_i] \geq \mathbb{E}[\pi_2|defect, q_i, p_i]$, so that $\rho_{21}(q_i, p_i) \geq 0$ and it is increasing in q_i. Hence, the more frequent cooperation is in the population expected by i, the stronger her indignation against defectors. Essentially, player i's indignation in the role of player 2 against a defecting player j measures the extent to which i considers player j responsible for i's expected losses compared to the counterfactual outcome based on i's beliefs (q_i, p_i) about the population play.

Let $\mathbb{E}[\pi_1|(defect, y_i), q_i, p_i]$ denote player i's expectation about the defector j's payoffs if i takes action $y_i \in \{0, 1\}$ in player role 2, and let $\mathbb{E}[\pi_2|(defect, y_i), q_i, p_i]$ denote player i's own expected payoffs in player role 2 given action y_i. I assume that individuals in player role 2 aim to maximize

$$u_i(y_i|\text{defect}, q_i, p_i) = \mathbb{E}[\pi_2|(defect, y_i), q_i, p_i] - \lambda_i\rho_{21}(q_i, p_i)\mathbb{E}[\pi_1|(defect, y_i), q_i, p_i],$$

[7] Not allowing player 2 to punish a cooperator is a simplification rather than a restriction. Regardless of individuals' beliefs about the average behavior of player 1 in the population, player 1's cooperation would never intend to harm player 2. Thus, player 2 would not feel indignation toward or punish a cooperator, although there might be reasons for doing so not captured in this paper (e.g., antisocial punishment documented in Herrmann et al. (2008)); see remarks in Section I, Online Appendix.

where $\lambda_i \geq 0$ measures i's primitive psychological tendency for indignation.[8] Since heterogeneity in this psychological parameter is not a focus of this paper, I assume $\lambda_i = \lambda \geq 0$ for each $i \in N$ to simplify the exposition.

Remarks. Indignation is triggered not by perceived harm alone but by a combination of perceived harm and a perceived violation of social norms. This norm-based aspect distinguishes it from Rabin (1993) and Dufwenberg and Kirchsteiger (2004)'s models of reciprocity, which ignore the broader social context and define kindness and unkindness based solely on beliefs regarding the strategies and beliefs of immediate interaction partners. Battigalli and Dufwenberg (2007)'s model of guilt and shame can motivate adherence to social norms but cannot explain the punishment of norm violators.

In Battigalli et al. (2019)'s model of frustration and anger, the beliefs that trigger anger are those about the behavior of immediate interaction partners. In contrast, the beliefs that trigger indignation against interaction partners are not about what I expect they *will* do, and their not doing so reduce my payoffs personally. Rather, indignation stems from what I expect they *should* do, based on what I anticipate most people would do in similar situations. A behavior feels wrong not only because it harms me and possibly others but, more importantly, because it also violates my perception of the social norm (Elster, 1998).

In addition to the conceptual differences, the departure from Battigalli et al. (2019) has behavioral consequences. For example, according to Battigalli et al. (2019), frustration and anger never occur on the equilibrium path (see their Proposition 1 for details). In contrast, a game can be constructed where a population of heterogeneous individuals are randomly matched to play a base game, and some individuals may consistently choose to act differently from the population average, triggering indignation by others in equilibrium.

2.2. Static equilibrium

Roughly speaking, an indignation equilibrium is a sequential equilibrium with decision utilities $u_i(x_i|\emptyset, q_i, p_i)$ and $u_i(y_i|\text{defect}, q_i, p_i)$. In general, individuals may be embedded in a neighborhood structure that gives them access to information from, and allows them to respond to, only a local subset of individuals. To accommodate this possibility, let the nonempty subset of individuals $M_i \subset N$ represent i's neighborhood. The case of $M_i = N$ for each i represents global interactions, where each individual may encounter any other in the population.

Definition 1. An *indignation equilibrium* is a state $z = (x, y, q, p)$ such that for each $i \in N$:

1. $q_i = \frac{1}{|M_i|}\sum_{j\in M_i} x_j$ and $p_i = \frac{1}{|M_i|}\sum_{j\in M_i} y_j$,
2. $x_i \in \{0,1\}$ maximizes $u_i(x_i|\emptyset, q_i, p_i)$, and
3. $y_i \in \{0,1\}$ maximizes $u_i(y_i|defect, q_i, p_i)$.

If $\lambda = 0$, then no one would ever punish defectors, and everyone defects. However, if λ is sufficiently large and many individuals expect others to cooperate, then many would punish defectors, thereby inducing cooperation and justifying their beliefs. Thus, a cooperation indignation equilibrium can arise.

The conditions for cooperation and punishment are as follows. Initially, there is no indignation present. Therefore, whether to cooperate or defect depends on which action maximizes expected material payoffs. Since $\mathbb{E}[\pi_1|(\emptyset, cooperate), q_i, p_i] = b$ and $\mathbb{E}[\pi_1|(\emptyset, defect), q_i, p_i] = (1-p_i)\bar{b}$, individual $i \in N$ weakly prefers to cooperate if and only if

$$p_i \geq \frac{\bar{b}-b}{\bar{b}} \equiv \gamma \qquad \text{(condition to cooperate)},$$

with $\gamma \in (0,1)$.

If player 1 defects, then in the role of player 2, i's indignation toward player 1 is

$$\begin{aligned}\rho_{21}(q_i, p_i) &= \mathbb{E}[\pi_2|\emptyset, q_i, p_i] - \mathbb{E}[\pi_2|defect, q_i, p_i] \\ &= [q_i\bar{d} + (1-q_i)(1-p_i)d] - (1-p_i)d \\ &= q_i(\bar{d} + p_i d - d).\end{aligned}$$

Note that $\rho_{21}(q_i, p_i) > 0$ for each $q_i > 0$, and $\rho_{21}(q_i, p_i) = 0$ if and only if $q_i = 0$. Thus, $u_i(y_i = 0|defect, q_i, p_i) = d - \lambda q_i(\bar{d} + p_i d - d)\bar{b}$, and $u_i(y_i = 1|defect, q_i, p_i) = 0$. Hence, i (weakly) prefers punishment if

$$\lambda q_i(\bar{d} + p_i d - d) \geq \frac{d}{\bar{b}}. \tag{1}$$

Therefore, whether i punishes defectors depends on both the expectation of the fraction of cooperators q_i and the expectation of the fraction of punishers p_i. While these expectations need not be equal, the case where $q_i = p_i$ will prove particularly useful in characterizing the long-run stable state considered later.[9]

[8] This specification follows Battigalli et al. (2019)'s modelling of "simple anger," where frustration increases with the shortfall between expected and realized payoffs. The assumption is consistent with classical findings from ultimatum game experiments where offers giving less of the pie are rejected more often, although these findings are usually interpreted as evidence for inequality aversion and reciprocity. Further, recent studies show that participants increase punishment of defectors when they expect more cooperation (Carpenter and Matthews, 2012; Li et al., 2021).

[9] In transitions from a defection state to a cooperation state, the least difficult path requires a sufficient number of players to switch to cooperation and punishment simultaneously. This simultaneous change then sets both belief elements, q_i and p_i, to the same value for all updating individuals.

Thus, let us consider the binding condition to induce punishment when $\lambda > 0$ and $q_i = p_i$. In this case, the minimum threshold for cooperators and punishers to induce punishment is given by

$$\eta \equiv \min\{q \mid \lambda q(\bar{d} + qd - d) \geq \frac{d}{b}\},$$

which is strictly positive, decreases with λ, and approaches zero as $\lambda \to \infty$. This sufficient condition to induce punishment can be succinctly expressed as

$$q_i = p_i > \eta \qquad \text{(condition to punish)}.$$

If $q_i = p_i < \eta$, then i does not punish.

Let $z^D \in Z$ denote the *defection state* such that all individuals defect, no one punishes defectors, and all individuals expect others to do the same. Let $z^C \in Z$ denote the opposite *cooperation state* such that all individuals cooperate, everyone punishes defectors if any exist, and everyone expects others to do the same. The following proposition shows that these two states can emerge as indignation equilibria. Although an inhomogeneous equilibrium, where some individuals behave differently from others, may exist under local interactions (when $M_i \subsetneq N$ for each i), these two homogeneous equilibria are the primary candidates for the long-run stable equilibrium considered later under both global and local interactions.[10]

Proposition 1. *Suppose $M_i = N$ for each $i \in N$. Then*

1. *the state z^D is an indignation equilibrium; and*
2. *if $\lambda > 0$ and $\eta < 1$ for each $i \in N$, then the state z^C is also an indignation equilibrium.*

The proof for Proposition 1 is straightforward given the above conditions for cooperation and punishment, so it is omitted. I define a *social norm* as a belief profile β contained in an indignation equilibrium (s, β). It is a set of beliefs such that, once established, everyone takes actions consistent with them, and their actions justify their beliefs. Proposition 1 implies that when $\lambda > 0$ and $\eta < 1$, there are two social norms that can be established. One supports cooperation and motivates people to punish defectors, while the other is associated with defection, under which defectors are not punished. Both are self-fulfilling. Throughout the rest of this paper, I assume $0 < \eta < 1$ to focus on the cases where multiple equilibria exist.

3. Dynamics

Building upon Kandori et al. (1993), Young (1993, 1998), and Ellison (1993, 2000), I consider an *unperturbed dynamic* in which each individual plays a rational strategy based on their beliefs each period, and these beliefs are formed by sampling the population's strategy profile from the previous period. I also consider a *perturbed dynamic* in which individuals occasionally make mistakes.

Time is discrete: $t = 0, 1, 2, \ldots$. The dynamics are Markov chains defined on the state space Z. The unperturbed dynamic is represented by the transition probability function $P^0 : Z \times Z \to [0, 1]$, where $P^0(z', z)$ denotes the probability of transitioning from state $z \in Z$ to state $z' \in Z$ within a period. An advantage of the Kandori et al. (1993) and Young (1993) approach is that it is not necessary to explicitly construct P^0. It suffices to specify the transitional events that occur with positive probability.[11] Each period consists of two stages of events. Let $z = (x, y, q, p)$ denote the population in the previous period.

Stage 1. Each $i \in N$ may revise one or both sides of their beliefs (q_i or p_i or both) with positive probability, and with complementary probability, i inherits beliefs from the last period.[12] If i is given the opportunity to revise beliefs, they randomly draw an *accessible sample* $M_i \subset N$ of size $0 < m \leq n$ from the population and calculate the fractions of cooperators and punishers based on their previous strategies. Depending on the collection of accessible samples $\{M_i\}$ that are drawn with positive probability for each individual, I distinguish between *global interactions* and *local interactions*.[13] While other local interaction structures are conceivable, I focus on the circular city model that has been widely used in the literature.[14]

Global interactions. The collection of accessible samples $\{M_i\}$ for each i consists of all non-empty subsets of N with size $|M_i| = m$.[15]

[10] As previous studies of stochastic stability analysis demonstrate, it is generally difficult to sustain inhomogeneous states in the long run (Young, 1993; Young and Burke, 2001; Kandori et al., 1993; Ellison, 2000), unless players choose their strategies and interaction neighbors simultaneously, i.e., when the interaction structure is endogenous (e.g., Cui and Weidenholzer, 2021), or when players have heterogeneous preferences.

[11] I thank an anonymous referee for suggesting the omission of the transition probability formulas to simplify the exposition.

[12] The possibility of inheriting beliefs captures a form of inertia and adjustment costs as assumed in Kandori et al. (1993) and Eshel et al. (1998).

[13] When multiple accessible samples can be drawn, it does not matter whether the sample is drawn uniformly from $\{M_i\}$ or according to some other distribution. As in Young (1993, 1998), it suffices that the drawing distribution has full support on the collection of accessible samples and that the probability does not vanish as $\varepsilon \to 0$.

[14] The circular city model is common in the literature, e.g., Dixit (2007), Ellison (1993), Bergstrom and Stark (1993), Eshel et al. (1998), Tabellini (2008). An earlier version of this paper, Chapter 2 of my PhD thesis, considers a more complex spatial structure in which individuals locate at the vertices of a two-dimensional lattice placed on a torus (Li, 2018), as examined by Ellison (2000). Despite the much longer proofs to deal with the intermediate evolution steps, the results are qualitatively identical: changing from global to local interactions does not change the long-run stable state.

[15] This sampling feature follows Young (1993) and Durieu and Solal (2003), which adds a bit of realism to the analysis and shows the results' generality. However, it is not essential; it can be assumed $m = n$. Under global interactions, the case of $m = n$ extends the learning model of Kandori et al. (1993).

Local interactions. Individuals $1, 2, \ldots, n$ are arranged in order on a circle, with individual n connected to individual 1 at one end. For each individual i, the accessible sample M_i includes i and $\frac{m-1}{2}$ individuals on each side of i.[16]

Given state $z = (x, y, q, p)$ in the last period, individual i updates to beliefs (q_i', p_i') with positive probability if and only if: (i) $q_i' = q_i$ and $p_i' = p_i$ in the event of inheriting beliefs, or (ii) in the event of revising beliefs, if $q_i' \neq q_i$, then there is an accessible sample $M_i \subset N$ for i such that $q_i' = \frac{1}{m} \sum_{j \in M_i} x_j$, and if $p_i' \neq p_i$, then there is an accessible sample $M_i \subset N$ with $p_i' = \frac{1}{m} \sum_{j \in M_i} y_j$. That is, individuals can sometimes revise only one side of their beliefs - either q_i or p_i - instead of always revising them simultaneously.[17] This captures the idea that, with some probability, an individual may primarily act in one player role within a certain duration of time and thus only observe a sample distribution of the actions of the opposite player role. Belief updating occurs independently across individuals and over time.

Stage 2. Each individual adopts a sequentially rational strategy based on their beliefs in the current period. More precisely, given each individual's belief (q_i', p_i') in the current period, they choose $x_i' \in \{0, 1\}$ and $y_i' \in \{0, 1\}$ to maximize $u_i(x_i'|\varnothing, q_i', p_i')$ and $u_i(y_i'|\text{defect}, q_i', p_i')$, respectively. The beliefs and strategy profiles $z' = (x', y', q', p')$ then become the current state. This completes the specification of P^0.

The perturbed dynamic P^ε differs from P^0 in that for each period, each $i \in N$ makes a mistake with probability $\varepsilon > 0$, in which case each $(x_i', y_i') \in \{(0,0), (1,0), (0,1), (1,1)\}$ is chosen with positive probability regardless of beliefs.

Let $\mu^\varepsilon : Z \to [0, 1]$ denote the stationary distribution of P^ε with $\varepsilon > 0$. According to the standard result that any finite irreducible Markov chain has a unique stationary distribution, μ^ε exists and is unique.[18] Following Young (1993), Ellison (1993), Young (1998), Ellison (2000), Kandori et al. (1993), Foster and Young (1990) and Ely (2002), I define stochastic stability for the game of cooperation and punishment under indignation:

Stochastic stability. A state $z \in Z$ is *long-run stable*, or *stochastically stable*, if

$$\lim_{\varepsilon \to 0} \mu^\varepsilon(z) > 0.$$

That is, a state is called a long-run stable equilibrium if it occurs with a positive probability in the stationary distribution of the perturbed dynamic as the probability of making mistakes vanishes. As the stationary distribution describes each state's empirical relative frequency in an infinite horizon, a long-run stable equilibrium must occur sufficiently frequently relative to all other states.

4. Global interactions

The perturbed dynamic P^ε approaches the unperturbed dynamic P^0 as $\varepsilon \to 0$, but they exhibit a crucial difference. The perturbed dynamic is an irreducible Markov chain: for any two states $z, z' \in Z$, there is a positive probability of transitioning from z to z' within a finite number of periods. In contrast, P^0 can be trapped in a state and stay there forever. An *absorbing set* is a minimal set of states within which a dynamic can be trapped. A state contained in a singleton absorbing set is called an *absorbing state*. I show in the appendix, Lemma 2, that there are only two absorbing sets of P^0, and they are singleton: one consists of the cooperation equilibrium z^C, and the other consists of the defection equilibrium z^D. Based on the theorems by Young (1993) and Ellison (2000), a long-run stable equilibrium can only be found in an absorbing set. Hence, z^C and z^D are the only candidates for a long-run stable equilibrium.

Which equilibrium–the cooperation or defection one–is long-run stable depends on the relative sizes of their *basins of attraction*. The basin of attraction of a state $z \in Z$, denoted by $B(z) \subset Z$, encompasses all states from which P^0 converges to z with positive probability. Once the unperturbed dynamic reaches $B(z^C)$ or $B(z^D)$, it converges to z^C or z^D with positive probability and remains there indefinitely in the absence of mistakes. Thus, once z^C or z^D is reached, transitions to another equilibrium will occur only when a sufficient number of individuals make a certain sequence of mistakes. For a small ε, it is extremely unlikely that multiple mistakes occur simultaneously. Therefore, the long-run equilibrium is determined by the minimum number of mistakes necessary to escape from $B(z^C)$ and $B(z^D)$, respectively.

Let $R(z)$ denote the minimum number of mistakes needed to leave $B(z)$ when starting from $z \in Z$, which is called the *radius* of $B(z)$. Define $\hat{\gamma}(q) = \min\{p | \lambda q(\bar{d} + pd - d) \geq \frac{d}{b}\}$, which is the minimum fraction of punishers p required for an individual to punish defectors when the fraction of cooperators is believed to be $q \in [0, 1]$. Let $\gamma^* = \max\{\gamma, \hat{\gamma}(1)\}$. The thresholds γ^* and η are critical in characterizing $R(z^D)$. They are the minimum fractions of punishers and cooperators, respectively, required to trigger a transition from the defection equilibrium to the cooperation equilibrium without further mistakes. Their complementary values, $1 - \gamma^*$ and $1 - \eta$, are the minimum fractions of non-punishers and defectors needed to trigger transitioning from the cooperation equilibrium to the defection equilibrium without further mistakes. Thus, they characterize $R(z^C)$. Let $\lceil r \rceil$ denote the smallest integer greater than or equal to number r.

[16] For the ease of exposition and without much loss of generality, I assume m to be an odd number under local interactions.

[17] The assumption that players may update only one component of their beliefs (either q_i or p_i) matters for how transition costs between equilibria are calculated along specific paths. However, it does not affect which equilibrium is stochastically stable. This is because requiring simultaneous updating of both belief components would increase the difficulty of leaving both $B(z^D)$ and $B(z^C)$ by the same proportion, leaving their relative transition costs-and thus stochastic stability-unchanged.

[18] Although beliefs q_i and p_i could take infinitely many values, what matters in the dynamics are the proportions of cooperators and punishers that can arise and be observed in the finite population, which take only finite possibilities. Hence, there is a finite number of states that the dynamics can transit to except for the initial states.

Lemma 1. *Consider global interactions. Then* $R(z^D) = \lceil m \min\{\eta, \gamma^*\} \rceil$ *and* $R(z^C) = \lceil m \min\{1-\eta, 1-\gamma^*\} \rceil$.

According to Young (1993)'s theorem, when there are only two absorbing states of P^0, the relative sizes of their radii fully characterize the stochastically stable state. Lemma 2 in the Appendix establishes that $\{z^C\}$ and $\{z^D\}$ are the only absorbing sets of P^0. Thus, Lemma 1 leads to the following result.

Theorem 1. *Consider global interactions. Suppose* $\lceil m\gamma^* \rceil \neq \lceil m(1-\eta) \rceil$ *and* $\lceil m\eta \rceil \neq \lceil m(1-\gamma^*) \rceil$.

1. *If* $\eta + \gamma^* > 1$, *then* z^D *is uniquely long-run stable.*
2. *If* $\eta + \gamma^* < 1$, *then* z^C *is uniquely long-run stable.*

When comparing the relative sizes of $R(z^C)$ and $R(z^D)$, the sample size m is canceled out, and the outcome depends solely on the thresholds γ^* and η. I offer several remarks below.

First, Theorem 1 reveals how the thresholds η and γ^* jointly determine long-run stability. Lower values of η and $\hat{\gamma}(1)$ reduce the bar to induce punishment, while lower values of γ reduces the bar to induce cooperation. Thus, lower values of η and $\gamma^* = \max\{\gamma, \hat{\gamma}(1)\}$ increase the chance of the cooperation equilibrium being long-run stable. Moreover, the condition $\eta + \gamma^* < 1$ for the long-run stability of cooperation implies that the factors underlying η and γ^* function as *substitutes*: the cooperation equilibrium becomes uniquely long-run stable when their sum is sufficiently small, indicating that policies promoting cooperation can be equally effective whether they target η or γ^*.

In the stochastic stability literature, $\frac{1}{2}$-dominance (Morris et al., 1995) typically suffices for long-run stability (see Weidenholzer (2010) for a survey). In the current model, z^C is $\frac{1}{2}$-dominant if and only if $\eta < \frac{1}{2}$ *and* $\gamma < \frac{1}{2}$, while z^D is $\frac{1}{2}$-dominant if and only if $\eta > \frac{1}{2}$ *and* $\gamma > \frac{1}{2}$.[19] Thus, in contrast with the conditions identified by Theorem 1, $\frac{1}{2}$-dominance implies that the factors behind η and γ work as *complements* - they each need to be smaller than a half. Therefore, while $\frac{1}{2}$-dominance remains sufficient for long-run stability in this model, it is unnecessary and may suggest suboptimal policy directions for promoting cooperation.

Second, my model combines two elements: the psychology of indignation and the best-reply updating rule. Indignation transforms the original material sequential game into a psychological game with multiple equilibria. A previous influential approach to cultural evolution instead applies the conformity updating rule to the original material game (e.g., Boyd and Richerson, 1985; Henrich and Boyd, 1998, 2001). Conformity presumes that individuals tend to adopt more prevalent strategies, and if material incentives support a less common strategy, individuals face trade-offs between the two considerations. Under this conformity rule, the cooperation state z^C can be an equilibrium without indignation. However, it never attains long-run stability, regardless of the strength of conformity. To illustrate this, suppose $m = n$ and that half of the population punishes defectors while the other half abstains at period t, resulting in equal frequency for both actions. Then, conformity loses its selection power, and material incentives favor non-punishment. Consequently, with positive probability, all individuals switch to non-punishment at $t+1$, and all individuals defect and do not punish at $t+2$, leading to the defection state z^D. It follows that $R(z^C) < \lceil n/2 \rceil$ while $R(z^D) \geq \lceil n/2 \rceil$, regardless of the conformity strength. Hence, z^D is always uniquely long-run stable. This comparison with the conformity-based approach highlights the essential role of indignation in upholding cooperation in the long run.

Finally, I provide a comparative statics remark regarding the original exogenous parameters of the base game and the psychological sensitivity parameter λ_i.

Remark 1 (Comparative statics). The threshold γ^* increases with the temptation to defect, $\bar{b} - b$, while η decreases with λ_i and the loss resulting from player 1's defection, $\bar{d} - d$. Therefore, if the temptation to defect is small, the loss resulting from others' defection is large, and individuals are sensitive to indignation, then $\eta + \gamma^*$ tends to be small, thereby increasing the chance of the cooperation equilibrium being long-run stable.

5. Local interactions and mobility

5.1. Local interactions in fixed neighborhoods

When counting the minimum numbers of mistakes for transitions, thresholds are rounded upward to integers. Thus, introducing the following notation simplifies the presentation of results. Let $q^\dagger = \frac{1}{m}(m + 1 - \lceil m\gamma \rceil)$, $\gamma^\dagger = \max\{\gamma, \hat{\gamma}(q^\dagger)\} + \frac{1}{m}$, $q^\ddagger = \frac{1}{m}(m - 1 - \lfloor m\gamma \rfloor)$, and $\gamma^\ddagger = \min\{\gamma, \hat{\gamma}(q^\ddagger)\} - \frac{1}{m}$. For a sufficiently large m, $\hat{\gamma}(q^\dagger)$ is arbitrarily close to $\hat{\gamma}(q^\ddagger)$. Hence, although $\gamma^\ddagger < \gamma^\dagger$, the difference between them is bounded by $|\gamma - \hat{\gamma}(q^\dagger)|$ plus an arbitrarily small number for a sufficiently large m. Additionally, the difference between $\gamma^\dagger$ and $\gamma^* = \max\{\gamma, \hat{\gamma}(1)\}$ is bounded by $\hat{\gamma}(q^\dagger) - \hat{\gamma}(1)$ plus an arbitrarily small number for a sufficiently large m.

Theorem 2. *Consider local interactions. Suppose* $n \geq 2m$ *and that* $m\eta$, $m\gamma^\dagger$, *and* $m\gamma^\ddagger$ *are not integers.*

1. *If* $\eta + \gamma^\dagger \leq 1$, *then* z^C *is uniquely long-run stable.*
2. *If* $\eta + \gamma^\ddagger \geq 1$, *then* z^D *is uniquely long-run stable.*

[19] Following the stochastic learning dynamics literature, risk dominance ($\frac{1}{2}$-dominant) refers to *strict* risk dominance ($\frac{1}{2}$-dominant).

Note that the conditions for the two equilibria to achieve long-run stability are qualitatively identical and quantitatively comparable to those under global interactions. This finding aligns with prior research on stochastic learning dynamics, which typically shows that the best-reply rule selects the same risk-dominant or $\frac{1}{2}$-dominant equilibrium regardless of global or local interactions (e.g., Ellison, 1993, 2000).

In this model, the equivalence between global and local interactions in long-run outcomes stems from how indignation creates a complementarity between cooperation and punishment: people cooperate when they expect punishment of defection, and punish when they expect sufficient cooperation. The key insight is that this complementarity not only supports the cooperation state as an equilibrium but also creates a *strategically symmetric* position for it relative to the defection equilibrium. Consequently, despite the more complex structure, it leads to a situation similar to simple pure coordination games: long-run stability depends on the minimum necessary number of mistakes for transitions between equilibria, which ultimately depend on the tipping thresholds for changing behavior at the local base game level, independent of the interaction structure.[20]

However, regarding the issue of the evolution of cooperation (norms), the equivalence between global interactions and local interactions departs from what might be expected from an imitation-based evolutionary model (Nowak and May, 1992; Eshel et al., 1998; Abramson and Kuperman, 2001; Santos and Pacheco, 2005; Szabó and Fáth, 2007; Nowak et al., 2010; Ohtsuki et al., 2006). An imitation-based model assumes that individuals imitate strategies yielding the highest average material payoffs (interpreted as behaviors exhibiting higher fitness) among their neighbors, a process that mirrors biological evolution. Eshel et al. (1998) show that under imitation, individuals can overcome the prisoner's dilemma and achieve efficient cooperation outcomes in the long run in the circular local interaction structure. In contrast, this cooperative outcome would not arise under global interactions or when individuals have the freedom to choose their interaction partners. In the context of coordination games, researchers have found that the imitation rule often selects the payoff-dominant equilibrium rather than the risk-dominant one in suitable local interaction structures (Alós-Ferrer and Weidenholzer, 2006, 2008). Therefore, the imitation-based approach would predict that cooperation is more likely to emerge and persist under local interactions in a fixed neighborhood structure. In contrast, the similarity between Theorems 1 and 2 implies that, from the standpoint of cooperation sustained by indignation-driven punishment threats, there would not be systematic differences in cooperation levels between global and local interactions within fixed neighborhood structures.

5.2. Mobility

Mobility refers to the free movement between communities. The following model adapts from a standard islands location model widely used in the literature (e.g., Oechssler, 1997; Ely, 2002; Blume and Temzelides, 2003).

Consider a set of communities, $V = \{1, 2\}$, and a total population of $2n$ individuals, $N = \{1, \dots, 2n\}$.[21] Let $1 \leq m < n$. Each community includes $(n - m)$ ***fixed residents*** who cannot migrate to another community. In reality, this immobility may be due to ethical, political, economic, or religious constraints, or a lack of information. The remaining $2m$ individuals in N are ***mobile individuals*** who can freely choose which community to live in each period.[22] The identities of fixed residents and mobile individuals can either be fixed over time or randomly re-selected each period. The ratio $\frac{m}{n}$ captures the *degree of mobility*: the proportion of individuals who can choose where to live.[23]

Consider a period t. Let p_{iv} denote individual i's belief regarding the proportion of punishers and q_{iv} denote their belief regarding the proportion of cooperators in community $v \in V$. Individual i's cooperation decision is denoted by $x_i \in \{0, 1\}$, and their punishment decision is denoted by $y_i \in \{0, 1\}$. The *state* for this period is $z = ((x_i, y_i, (q_{iv}, p_{iv})_{v\in V})_{i\in N}, (N_v)_{v\in V})$, which describes each individual's decisions, their beliefs p_{iv} and q_{iv} for each community, and a partition of N into communities, $(N_v)_{v\in V}$. Let Z denote the collection of all states. How the population transitions from one state $z \in Z$ to another $z' \in Z$ is determined by the following events within each period:

Stage 1. The mobile individuals decide which community $v \in V$ to live in for the current period. To make this decision, they calculate the expected utilities for each community and choose the one with higher expected utilities, given the expectation that they will take a strategy to maximize expected utilities given their beliefs $(q_{iv}, p_{iv})_{v\in V}$ at the current period. If the expected utilities are the same for the two communities, then they randomly pick one.[24] The set $N_v \subset N$ includes the fixed residents of community v and the mobile individuals who choose to live in community v for the current period.

[20] However, the speed of convergence can differ significantly; convergence can be much faster in local interactions, as demonstrated by Ellison (1993, 2000). Moreover, the equivalence between global and local interactions in long-run outcomes should be interpreted with caution, as I only consider the circle network here (and a two-dimensional lattice in an earlier version of this paper). As shown in Section III of the Online Appendix, asymmetric structures like the star network can yield qualitatively different results, where z^C and z^D can simultaneously be long-run stable in a broad range of parameters. The bottom line here is that making interactions local per se does not necessarily lead to more cooperation.

[21] The setup and results can bed extended and apply to multiple communities and the chosen notation can be readily generalized.

[22] The groups of fixed residents and mobile individuals may change or be randomly re-selected each period without altering the argument or any of the results.

[23] Similar mobility models with finite neighborhoods have been previously explored by Ely (2002) and Bhaskar and Vega-Redondo (2004) who study coordination games with standard preferences.

[24] The probabilities of choosing the two communities at random can be heterogeneous across individuals. The only requirement is that the probabilities are invariant over time.

Stage 2. For the unperturbed dynamic P^0, each $i \in N$ chooses $x_i \in \{0,1\}$ and $y_i \in \{0,1\}$ to maximize $u_i(x_i|\emptyset, q_{iv}, p_{iv})$ and $u_i(y_i|\text{defect}, q_{iv}, p_{iv})$, respectively, for the community in which she lives, conditional on beliefs (q_{iv}, p_{iv}). For the perturbed dynamic P^ε, individuals make mistakes and randomly select a strategy with probability $\varepsilon > 0$.

Stage 3. Let q'_{iv} and p'_{iv} denote i's beliefs in the next period. For each $i \in N$ and $v \in V$, there are four complementary belief updating events, each occurring with a positive probability: (i) $q'_{iv} = q_{iv}$ and $p'_{iv} = p_{iv}$, (ii) $q'_{iv} = q_{iv}$ and $p'_{iv} = \frac{1}{|N_v|}\sum_{j\in N_v} y_j$, (iii) $q'_{iv} = \frac{1}{|N_v|}\sum_{j\in N_v} x_j$ and $p'_{iv} = p_{iv}$, and (iv) $q'_{iv} = \frac{1}{|N_v|}\sum_{j\in N_v} x_j$ and $p'_{iv} = \frac{1}{|N_v|}\sum_{j\in N_v} y_j$. All random events are independent across individuals and over time.

Let $\mathcal{D} \subset Z$ include all *defection states* where all individuals defect in each community, no individual punishes defectors, and everyone expects all individuals to do the same. That is, for each $z \in \mathcal{D}$, we have $x_i = y_i = 0$ and $p_{iv} = q_{iv} = 0$ for each $v \in V$ and $i \in N$. The states in $\mathcal{D}$ differ only in how the mobile individuals are distributed across communities. Similarly, let $\mathcal{C} \subset Z$ denote the set of *cooperation states* such that for each $z \in \mathcal{C}$, we have $x_i = y_i = 1$ and $p_{iv} = q_{iv} = 1$ for each $v \in V$ and $i \in N$. Relatedly, for each $v \in V$, we use $\mathcal{C}_v \subset Z$ to denote the subset of states such that (i) community v populates all mobile individuals so that its size is maximized, $n+m$, whereas community $v' \neq v$ contains only the fixed residents for the community so that its size is minimized, $n-m$; (ii) every $i \in N_v$ takes $x_i = y_i = 1$, whereas every $i \in N_{v'}$ takes $x_i = y_i = 0$; and (iii) every $i \in N$ believes $p_{iv} = q_{iv} = 1$ and $p_{iv'} = q_{iv'} = 0$. That is, the states in $\mathcal{C}_v$ represent a heterogeneous population where the cooperation norm prevails in community v while the defection norm prevails in the other community, and the size of community v is maximized.[25] The threshold γ^* is the same one defined for Theorem 1: $\gamma^* = \max\{\gamma, \hat{\gamma}(1)\}$.

Theorem 3. *Consider the model under mobility. Suppose that* $\lceil \gamma^*(n-m)\rceil \neq \lceil (1-\eta)(n-m)\rceil$, $\lceil \eta(n-m)\rceil \neq \lceil (1-\gamma^*)(n-m)\rceil$, *and* $b + \bar{d} > \bar{b} + d$.

1. *If* $\eta + \gamma^* < 1$, *then the long-run stable equilibria are the states in* $\mathcal{C}$.
2. *If* $1 < \eta + \gamma^* < 1 + \frac{m}{n}(1 - |\eta - \gamma^*|)$, *then the long-run stable equilibria are the states in* $\mathcal{C}_1 \cup \mathcal{C}_2$.
3. *If* $\eta + \gamma^* > 1 + \frac{m}{n}(1 - |\eta - \gamma^*|)$, *then the long-run stable equilibria are the states in* $\mathcal{D}$.

Observe the differences between this statement and Theorem 2 that deals with local interactions within a fixed neighborhood structure. The differences highlight how mobility can promote cooperation through two channels.

First, without mobility, the defection state is long-run stable only if $\eta + \gamma^\ddagger > 1$. With mobility, the mixture states in $\mathcal{C}_1 \cup \mathcal{C}_2$ are stable as long as $1 < \eta + \gamma^* < 1 + \frac{m}{n}(1 - |\eta - \gamma^*|)$. In these mixture states, the cooperation norm is established in the maximal community that consists of the community's fixed residents and all mobile residents in the population. Since $\gamma^\ddagger$ approximates γ^* for large population and sample sizes, this enlarges the range of parameters through which the cooperation norm can emerge and persist in the maximal community. The higher the degree of mobility, $\frac{m}{n}$, the wider the parameter range $\eta + \gamma^* \in (1, 1 + \frac{m}{n}(1 - |\eta - \gamma^*|))$ becomes to support the long-run stability of states in $\mathcal{C}_1 \cup \mathcal{C}_2$.

Second, mobility can promote cooperation by increasing the size of the cooperative communities in the states in $\mathcal{C}_1 \cup \mathcal{C}_2$. This is due to a "voting-with-feet" effect. Consider a community where the cooperation norm prevails and another where defection is the norm. Initially, the cooperative community might be smaller, but the condition $b + \bar{d} > \bar{b} + d$ implies that individuals anticipate higher payoffs in the cooperative community. Consequently, all mobile individuals gravitate toward the cooperative community. Individuals will *not* move to a cooperative community and then defect to exploit the cooperators because any residents with indignation are prepared to punish defectors within a cooperative community. Instead, they will adapt to the local cooperation norm and choose to cooperate.

Now, I explain how mobility can support the long-run stability of states in $\mathcal{C}_1 \cup \mathcal{C}_2$. First, assume that a community has locally established the cooperation norm while the defection norm prevails in another community. Then any states where the cooperative community is not maximal are transitional, as every mobile individual will opt to join the cooperative community for higher expected payoffs. Hence, in any absorbing sets, the cooperative community must reach its maximum size $n+m$, leading to a state in $\mathcal{C}_1 \cup \mathcal{C}_2$. Second, when $\eta + \gamma^* > 1$, the long-run stability of states in $\mathcal{C}_1 \cup \mathcal{C}_2$ depends on the minimum number of mistakes needed to transition from $\mathcal{C}_1 \cup \mathcal{C}_2$ to $\mathcal{D}$, compared to transitions in the opposite direction. It turns out that transitions from $\mathcal{C}_1 \cup \mathcal{C}_2$ to $\mathcal{D}$ are more difficult if $\eta + \gamma^* < 1 + \frac{m}{n}(1 - |\eta - \gamma^*|)$. This is because transitions to $\mathcal{D}$ require a sufficient number of individuals to defect or stop punishing by making mistakes in the *maximal* cooperative community whose size is $n+m$. Conversely, for transitions from $\mathcal{D}$ to $\mathcal{C}_1 \cup \mathcal{C}_2$, we can wait for the defection community to reach the *smallest* size $n-m$, which requires no mistakes, and then require a sufficient number of individuals to either cooperate or punish by mistake in this small defection community. Therefore, as long as $\eta + \gamma^*$ is not too large to offset this community size advantage of cooperation norms over defection norms, $\mathcal{C}_1 \cup \mathcal{C}_2$ will be more stable than $\mathcal{D}$.

The main driving force follows a standard group selection type of argument: groups practicing cooperative norms perform better, thereby attracting more residents and outcompeting other groups. This result aligns with previous findings from the stochastic stability literature (Blume and Temzelides, 2003; Anwar, 2002; Oechssler, 1999; Dieckmann, 1999; Ely, 2002). My departure is that prior work starts from simple coordination games and takes the existence of multiple, Pareto-ranked equilibria as granted, providing no explanation for why the original prisoner's dilemma game can give rise to multiple absorbing states analogous to the competition between risk-dominant versus payoff-dominant equilibria. My model provides a psychological foundation for that: when individuals move to a cooperative community, they do not conform to local behavior mechanically but internalize the local norm through updating their beliefs about what constitutes appropriate behavior in that society, which in turn triggers their indignation against defectors.

[25] If fixed residents and mobile individuals are fixed over time, then $\mathcal{C}_v$ is singleton.

This explains why mobile individuals could adapt to and reinforce local cooperation norms rather than exploit and undermine the cooperative atmosphere.

6. Conclusion

This study provides a microfoundation for social norms that sociology has long employed to explain behavior, and shows how individual emotions, a focus of psychology, can have enduring aggregate consequences. This is accomplished by combining psychological game theory (Geanakopolos et al., 1989; Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Battigalli and Dufwenberg, 2007, 2009; Battigalli et al., 2019) with stochastic stability analysis (Kandori et al., 1993; Ellison, 1993; Young, 1993; Ellison, 2000; Foster and Young, 1990), and focusing on cooperation and punishment in social dilemmas. The model assumes that, at the individual level, people experience the belief-dependent emotion of indignation: they feel angry toward those who violate common social practices and harm others. Consequently, both cooperation and defection can be sustained as self-fulfilling social norms. I then examine the cooperation norm's stability in a stochastic dynamic.

Two results emerge from the analysis. First, contrary to the predictions from prior imitation-based evolutionary approach, cooperation levels under global interactions can be as high as those under local interactions within a fixed neighborhood structure. Second, mobility facilitates cooperative communities to grow, promoting both their emergence and persistence.

These results speak to a number of recent empirical findings. In their well-known cross-cultural study, Henrich et al. (2010) conduct a series of behavioral experiments across a diverse range of societies, from small-scale societies to large, modern ones. They discover that egalitarian actions, classified as cooperative actions here, are more prevalent in large, modern societies with higher degrees of market integration as compared to small-scale societies.[26] One way to interpret their findings is that large modern societies with greater market integration are associated with enhanced mobility. My theoretical results thus provide a potential explanation for their empirical findings. Indeed, several recent experimental studies find that local interactions in fixed neighborhoods do not significantly improve cooperation over global interactions (Grujic et al., 2010; Rand et al., 2011; Kirchkamp and Nagel, 2007). Cooperation levels, however, are found to be higher in dynamic networks where individuals can change their interaction partners (Rand et al., 2011).[27]

Data availability

No data was used for the research described in the article.

Declaration of competing interest

I hereby declare that there are no conflicts of interest to report in connection with my submission titled "Indignation and the Evolution of Cooperation Norms" to Games and Economic Behavior. This work is the result of my independent research and has not been influenced by any external agencies or personal relationships that could be construed as a potential conflict of interest.

Appendix A. Proofs

Proof of Lemma 1

Suppose the dynamic P^ε starts from z^D at $t = 0$. To leave $B(z^D)$, assume $k \le n$ individuals all make mistakes and choose $x_i = y_i = 1$ in a period, say $t = 1$. In period $t = 2$, assume all individuals receive the opportunity to revise beliefs, and each draws a sample that includes all of the k individuals with $x_i = y_i = 1$ from the previous period.

Case 1: Suppose $\eta \le \gamma^* = \max\{\gamma, \hat{\gamma}(1)\}$. Then, let $k \ge \lceil m\eta \rceil$, and let all individuals update their beliefs to $q_i = p_i = \frac{k}{m} \ge \eta$. As a best reply, all individuals start to punish defectors, taking $y_i = 1$. In period $t = 3$, let all individuals revise, and only revise, their belief about the proportion of punishers in the population, leading to $p_i = 1 > \gamma$ and $q_i = \frac{k}{m} \ge \eta$. Then all individuals take $x_i = y_i = 1$ as a best reply.

Case 2: Suppose $\gamma^* < \eta$. When $k \ge \lceil m\gamma^* \rceil$, and all individuals update their beliefs to $q_i = p_i = \frac{k}{m} \ge \gamma^* \ge \gamma$, they will cooperate as a best-reply by taking $x_i = 1$. Then at $t = 3$, when all individuals revise (and only revise) their belief about the proportion of cooperators in the population, resulting in $q_i = 1$ and $p_i = \frac{k}{m} \ge \gamma^* \ge \hat{\gamma}(1)$, it follows from the definition of $\hat{\gamma}(1)$ that all individuals will take $x_i = y_i = 1$ as a best response.

In either case, the dynamic escapes $B(z^D)$ and falls into $B(z^C)$ within three periods. If $k < \lceil m\gamma^* \rceil$ and $k < \lceil m\eta \rceil$, then neither of the above transition paths will occur, and, absent any further mistakes, the dynamic will converge back to z^D. Therefore, the minimum number of necessary mistakes for the transitions is $\lceil m \min\{\eta, \gamma^*\} \rceil$.

[26] Market integration in their study is measured by the percentage of food obtained from market transactions within a society.

[27] In Rand et al. (2011)'s experiment, participants play the prisoner's dilemma game on networks. There are three treatments: the random-link treatment (the network is randomly regenerated for each round), the fixed-link treatment (the network is fixed during the experiment), and the endogenous-link treatment (participants can choose to rewire links). They find that cooperation is on average not statistically different between the random-link treatment and the fixed-link treatment; in both treatments, cooperation decays over time. In contrast, cooperation persists at a high level in the endogenous-link treatment.

Conversely, suppose we start from z^C at $t = 0$. To leave $B(z^C)$, assume $k \leq n$ individuals all make mistakes and choose $x_i = y_i = 0$ in a period, say $t = 1$. In period $t = 2$, all individuals revise their beliefs, and each draws a sample that includes all of the k individuals with $x_i = y_i = 0$ from the previous period.

Case 1: $1 - \eta \leq 1 - \gamma^*$. Let $k \geq \lceil m(1-\eta) \rceil$, and all individuals update their beliefs to $q_i = p_i = \frac{m-k}{m} \leq \eta$. As a best reply, all $i \in N$ take $y_i = 0$. In period $t = 3$, let all revise and only revise their belief about the proportion of punishers in the population, leading to $p_i = 0 < \gamma$ and $q_i = \frac{m-k}{m} \leq \eta$. Then all individuals take $x_i = y_i = 0$. In the next period, let all individuals update beliefs; then the dynamic falls to z^D.

Case 2: $1 - \gamma^* < 1 - \eta$. Let $k \geq \lceil m(1-\gamma^*) \rceil$. At $t = 2$, all individuals update their beliefs to $q_i = p_i = \frac{m-k}{m} \leq \gamma^*$. If $\gamma^* = \gamma$, then $p_i \leq \gamma$ for each $i \in N$, and all individuals take $x_i = 0$ as a best response. At $t = 3$, all individuals revise their beliefs to $q_i = 0$ and $p_i = \frac{m-k}{m} \leq \gamma$, taking $x_i = y_i = 0$ as a best response. If $\gamma < \gamma^* = \hat{\gamma}(1)$, then at $t = 2$, we have $q_i < 1$ and $p_i \leq \hat{\gamma}(1)$, so we can require all individuals to take $y_i = 0$ as a best response. At $t = 3$, all individuals update their beliefs to $p_i = 0$ and $q_i = \frac{m-k}{m} \leq \hat{\gamma}(1)$, and best-reply with $x_i = y_i = 0$. By requiring all individuals to update their beliefs in the next period, the dynamic falls to z^D.

If $k < \lceil m(1-\eta) \rceil$ and $k < \lceil m(1-\gamma^*) \rceil$, then $q_i = p_i = \frac{m-k}{m} > \max\{\eta, \gamma^*\}$ at $t = 2$. The first transition path would not work without further mistakes. For the second path, since $p_i > \gamma^* \geq \gamma$, taking $x_i = 0$ at $t = 2$ is not a best response. Additionally, $q_i = p_i > \eta$ implies that taking $y_i = 0$ is not a best response. Therefore, we must have $x_i = y_i = 1$ for each individual at $t = 2$, ensuring that the dynamic remains in $B(z^C)$. Consequently, the minimum number of mistakes required to escape from $B(z^C)$ is $\lceil m \min\{1-\eta, 1-\gamma^*\} \rceil$. This completes the proof of Lemma 1.

Proof of Theorem 1

The proof of Theorem 1 applies Young's theorem (1993): Let Ω denote the collection of all absorbing sets of P^0. A stochastic tree of an absorbing set $L \in \Omega$ is a function $t : \Omega \to \Omega$ such that for each $L' \in \Omega, L' \neq L$, there is a natural number k with $t^k(L') = L$, where t^k is the k-th functional power of t. It represents a directed tree graph rooting at and pointing toward L. For any two absorbing sets L and L', let $c(L, L')$ be the minimum number of mistakes required to transition from L to L'. The cost of a stochastic tree t of L is defined and denoted by $c(t) = \sum_{L' \in \Omega, L' \neq L} c(L', t(L'))$.

Theorem (Young, 1993). *A state $z \in Z$ is stochastically stable if and only if it is contained in an absorbing set $L \in \Omega$ that minimizes the cost among all stochastic trees over all absorbing sets in Ω.*

According to Young's theorem, establishing Theorem 1 requires combining Lemma 1 with the following lemma, which shows that P^0 converges to either z^D or z^C.

Lemma 2. *Under global interactions, $\{z^D\}$ and $\{z^C\}$ are the only absorbing sets of P^0.*

Proof of Lemma 2. Suppose another absorbing set of P^0 exists that includes $z \in Z$. Then $z \notin B(z^D) \cup B(z^C)$, and there is no state $z' \in B(z^D) \cup B(z^C)$ such that P^0 can move from z to z' within a finite number of periods with positive probability. I will show counterexamples below for contradiction. Consider $z = (x, y, q, p)$ at $t = 0$. Let $M \subset N$ be a subset with an arbitrary sample of m individuals, and write $\hat{q} = \sum_{i \in M} x_i/m$ and $\hat{p} = \sum_{i \in M} y_i/m$.

First, suppose $\lambda \hat{q}(\bar{d} + \hat{p}d - d) \leq \frac{d}{b}$. Then, at $t = 1$, let all individuals update beliefs to $q_i = \hat{q}$ and $p_i = \hat{p}$ by drawing sample M. Next, let all individuals take $y_i = 0$ at $t = 1$, which is a best response and occurs with positive probability. At $t = 2$, let all individuals update their belief about the fraction of punishers to $p_i = 0$ with q_i unchanged. Then, all individuals take $x_i = y_i = 0$ as a best response, leading to a state $z' \in B(z^D)$ with positive probability.

Second, suppose $\lambda \hat{q}(\bar{d} + \hat{p}d - d) > \frac{d}{b}$. At $t = 1$, let all individuals update beliefs to $q_i = \hat{q}$ and $p_i = \hat{p}$ by drawing sample M, and all take $y_i = 1$ as a best response. Then at $t = 2$, all individuals update their belief to $p_i = 1$ while q_i remains unchanged. All individuals will then take $x_i = y_i = 1$ as a best response, leading to a state $z' \in B(z^C)$. $\square$

According to Lemma 2, s^C and s^D are the only candidates for the long-run stable equilibrium. Therefore, comparing $R(s^C)$ with $R(s^D)$ fully characterizes the long-run stable equilibrium. Given Lemma 1, what remains to be shown is that if $\lceil \gamma^* m \rceil \neq \lceil (1-\eta)m \rceil$ and $\lceil \eta m \rceil \neq \lceil (1-\gamma^*)m \rceil$, then $\eta + \gamma^* > 1$ implies $\lceil \min\{\eta, \gamma^*\}m \rceil > \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$, and $\eta + \gamma^* < 1$ implies $\lceil \min\{\eta, \gamma^*\}m \rceil < \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$. Below, I consider the cases of $\eta \geq \gamma^*$ and $\eta < \gamma^*$ respectively.

First, suppose $\eta \geq \gamma^*$. Then $\min\{\eta, \gamma^*\} = \gamma^*$ and $\min\{1-\eta, 1-\gamma^*\} = 1-\eta$. If $\eta + \gamma^* > 1$, then $\gamma^* m > (1-\eta)m$, and thus $\lceil \min\{\eta, \gamma^*\}m \rceil = \lceil \gamma^* m \rceil \geq \lceil (1-\eta)m \rceil = \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$. Since $\lceil \gamma^* m \rceil \neq \lceil (1-\eta)m \rceil$, we obtain $\lceil \min\{\eta, \gamma^*\}m \rceil > \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$. Conversely, if $\eta + \gamma^* < 1$, then $\gamma^* m < (1-\eta)m$. Therefore, given $\lceil \gamma^* m \rceil \neq \lceil (1-\eta)m \rceil$, we obtain $\lceil \min\{\eta, \gamma^*\}m \rceil = \lceil \gamma^* m \rceil < \lceil (1-\eta)m \rceil = \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$.

Second, suppose $\eta < \gamma^*$. Then $\min\{\eta, \gamma^*\} = \eta$ and $\min\{1-\eta, 1-\gamma^*\} = 1-\gamma^*$. If $\eta + \gamma^* > 1$, then $\eta m > (1-\gamma^*)m$. Thus, given $\lceil \eta m \rceil \neq \lceil (1-\gamma^*)m \rceil$, we obtain $\lceil \min\{\eta, \gamma^*\}m \rceil = \lceil \eta m \rceil > \lceil (1-\gamma^*)m \rceil = \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$. If $\eta + \gamma^* < 1$, then $\eta m < (1-\gamma^*)m$, and consequently, $\lceil \min\{\eta, \gamma^*\}m \rceil = \lceil \eta m \rceil < \lceil (1-\gamma^*)m \rceil = \lceil \min\{1-\eta, 1-\gamma^*\}m \rceil$. This completes the proof for Theorem 1.

Proof of Theorem 2

The co-radius of a state $z \in Z$, denoted by $CR(z)$, is the minimum number of mistakes necessary to reach z when starting from the most distant state among all states. Let $\tau(z, z')$ be the minimum number of mistakes required to transition from z to z'. Then $CR(z) = \max_{z' \in Z} \min_{(z_1, z_2, \ldots, z_k)} \sum_{i=1}^{k-1} \tau(z_i, z_{i+1})$, where $(z_1, z_2, \ldots, z_k)$ is a transition path from $z_1 = z'$ to $z_k = z$ and the minimum is taken over all paths from z' to z. Further, for an absorbing set L of P^0, let $R(L)$ denote the minimum number of mistakes needed to leave the basin of attraction of L when the dynamic starts from a state in L. Consider a path $(z_1, z_2, \ldots, z_k)$ that goes through absorbing sets $L_1, L_2, \ldots, L_\kappa$ where, if z_1 is contained in an absorbing set, then L_1 is the absorbing set containing z_1, and if z_k is contained in an absorbing set, then L_κ is the absorbing set containing z_k. We have $\kappa \le k$ and that κ may be strictly smaller than k.

Following Ellison (2000), I define the modified cost of the path by $\tau^*(z_1, z_2, \ldots, z_k) = \sum_{i=1}^{k-1} \tau(z_i, z_{i+1}) - \sum_{i=2}^{\kappa-1} R(L_i)$, and the modified co-radius of $z \in Z$ by

$$CR^*(z) = \max_{z' \in Z} \min_{(z_1, z_2, \ldots, z_k)} \tau^*(z_1, z_2, \ldots, z_k)$$

where the minimum is taken over all paths $(z_1, z_2, \ldots, z_k)$ from $z_1 = z'$ to $z_k = z$. Ellison (2000)'s Theorems 1 and 3, combined, imply:

Theorem (Ellison, 2000). *If $R(z) > CR^*(z)$, then z is uniquely stochastically stable.*

I use $M_i \subset N$ with $|M_i| = m$ to denote individual i's neighborhood, which includes the $\frac{m-1}{2}$ individuals on the left of i on the circle, the $\frac{m-1}{2}$ individuals on the right, and i herself. Theorem 2 is established through four claims. By Ellison's theorem, Claims 1 and 4 together imply $R(z^C) > CR^*(z^C)$ given the condition in the first bullet point of Theorem 2; thus, z^C is uniquely stochastically stable. By contrast, Claims 2 and 3 together imply $R(z^D) > CR^*(z^D)$ given the condition in the second bullet point of Theorem 2, so z^D is uniquely stochastically stable.

Claim 1. *Let $q^\dagger = (m + 1 - \lceil m\gamma \rceil)/m$ and $\gamma^\dagger = \max\{\hat{\gamma}(q^\dagger), \gamma\}$. Suppose $m\eta$ and $m\gamma^\dagger$ are not integers. If $\eta + \gamma^\dagger < 1$, then $CR^*(z^C) \le R(z^D) = \lceil m \min\{\eta, \gamma^\dagger\} \rceil$.*

Proof of Claim 1. Suppose that the initial state is z^D, and consider how to escape $B(z^D)$. Let $J = \{1, 2, \ldots, k\} \subset N$ denote the set of individuals, with $1 \le k \le n$, who make mistakes; each j in J is required to switch to $x_j = y_j = 1$ by mistake. Denote $k^* = \min\{k, \frac{m+1}{2}\}$. Assume $\eta + \gamma^\dagger < 1$. There are two cases: $\eta \le \gamma^\dagger$ and $\eta > \gamma^\dagger$.

Case 1: Suppose $\eta \le \gamma^\dagger$. Then $\eta + \gamma^\dagger < 1$ implies $\eta < \frac{1}{2} < \frac{m+1}{2m}$. Suppose that $k > m\eta$ holds. Then $\frac{k^*}{m} > \eta$. Denote $J' = \{i \in N | \frac{|J \cap M_i|}{m} > \eta\}$. Let each $i \in J'$ update beliefs to $q_i = p_i = \frac{|J \cap M_i|}{m} > \eta$ and take $y_i = 1$ as the unique best-reply. The condition $\frac{k^*}{m} > \eta$ implies $J \subset J'$ and that the size of $J' \backslash J$ is $m + 1 - 2\lceil m\eta \rceil \ge 0$. Next, consider individual k, who is located at the boundary of J. Note that $|J' \cap M_k| = m + 1 - \lceil m\eta \rceil$, since J' includes $\frac{m+1}{2}$ individuals on the left-hand side of k (plus k themselves) and $\frac{m+1}{2} - \lceil m\eta \rceil$ individuals on the right-hand side of k. If $\frac{|J' \cap M_k|}{m} = \frac{m+1-\lceil m\eta \rceil}{m} > \gamma$ (condition $**$) holds, then let k and all other individuals in J update their beliefs, resulting in $p_j > \gamma$, $p_j > \eta$, and $q_j > \eta$ for each $j \in J$, who best-responds with $x_j = y_j = 1$. The dynamic then reaches a state where the k individuals in J are cooperating and punishing, and all individuals in J' (which contains J) are punishing. Without mistakes, no one in J would change their behavior, and no one in J' will change their punishment decisions in any future period. Thus, the dynamic has escaped $B(z^D)$. Hence, a sufficient condition to escape $B(z^D)$ is: (i) $\frac{m+1}{2} > m\eta$ and $k > m\eta$ so that $k^* = \min\{k, \frac{m+1}{2}\} > m\eta$, and (ii) $m + 1 - \lceil m\eta \rceil > m\gamma$. The latter is implied by $\eta + \gamma \le \eta + \gamma^\dagger < 1$, because $m + 1 - \lceil m\eta \rceil \ge m - m\eta = m(1 - \eta) > m\gamma$. The condition $k > m\eta$ is also necessary; otherwise, $J' \subset J$, and both J and J' would shrink over time with positive probability and converge back to z^D with probability one. Hence, the minimum number of mistakes k for the transition is $\lceil m\eta \rceil$ given $\eta < \frac{m+1}{2m}$, $\eta + \gamma^\dagger < 1$, and $m\eta$ not an integer.

Case 2: Suppose $\eta > \gamma^\dagger$. Then $\eta + \gamma^\dagger < 1$ implies $\gamma \le \gamma^\dagger < \frac{1}{2} < \frac{m+1}{2m}$. Suppose $k^* > m\gamma$. Denote $J'' = \{i \in N | \frac{|J \cap M_i|}{m} > \gamma\}$. Let each $i \in J''$ update beliefs to $q_i = p_i = \frac{|J \cap M_i|}{m} > \gamma$ and take $x_i = 1$ as a best reply. The condition $\frac{k^*}{m} > \gamma$ implies $J \subset J''$ and that $J'' \backslash J$ includes $\frac{m+1}{2} - \lceil m\gamma \rceil \ge 0$ individuals on each of the two sides of J. Consider individual k. We have $|J'' \cap M_k| = m + 1 - \lceil m\gamma \rceil$ and $q^\dagger \equiv \frac{|J'' \cap M_k|}{m} = \frac{1}{m}(m + 1 - \lceil m\gamma \rceil)$, because J'' includes $\frac{m+1}{2}$ individuals on the left-hand side of k (plus k herself) and $\frac{m+1}{2} - \lceil m\gamma \rceil$ individuals on the right-hand side of k. Now, let k and each $j \in J$ update their beliefs, resulting in $q_j \ge \frac{|J'' \cap M_k|}{m} = q^\dagger$ and $p_j \ge \frac{k^*}{m} > \hat{\gamma}(q^\dagger)$ for each $j \in J$. The conditions $q_j \ge q^\dagger$, $p_j > \gamma$, and $p_j > \hat{\gamma}(q^\dagger)$ together imply $x_j = 1$ and $y_j = 1$ as the unique best reply. Without further mistakes, no one in J and J'' would change their actions. We thus have escaped $B(z^D)$. The initial assumption $k^* > m\gamma^\dagger$, i.e., $\frac{m+1}{2} > m\gamma^\dagger$ and $k > m\gamma^\dagger$, is both sufficient and necessary. Hence, the minimum number of mistakes k to escape $B(z^D)$ is $\lceil m\gamma^\dagger \rceil$, given $\gamma^\dagger < \frac{1}{2} < \frac{m+1}{2m}$ and $m\gamma^\dagger$ not an integer.

Combining Case 1 and Case 2, we obtain $R(z^D) = \lceil m \min\{\eta, \gamma^\dagger\} \rceil$.

Next, to determine $CR^*(z^C)$, consider transitions from the most distant state z^D to z^C. First, observe that the sets J, J', J'' above depend on k, so we can index them by k: J_k, J'_k, J''_k. Individual $k' = |J'_k|$ is at the boundary of J'_k, and individual $k'' = |J''_k|$ is at the boundary of J''_k. Second, we have $|J_{k+1} \cap M_{k'+1}| = |J_k \cap M_{k'+1}| - 1 = |J_k \cap M_{k'}| + 1 - 1 = |J_k \cap M_{k'}|$. Hence, for Case 1, $\frac{|J_k \cap M_{k'}|}{m} > \eta$ implies $\frac{|J_{k+1} \cap M_{k'+1}|}{m} > \eta$. Thus, one extra mutation of individual $k + 1$ to switch to $x_{k+1} = y_{k+1} = 1$ is sufficient to expand the sets of cooperators and punishers to J_{k+1} and J'_{k+1}. Similarly, for Case 2, $\frac{|J_k \cap M_{k''}|}{m} > \gamma$ implies $\frac{|J_{k+1} \cap M_{k''+1}|}{m} > \gamma$, and therefore it takes at most one extra mutation to expand the sets of cooperators and punishers to J_{k+1} and J''_{k+1}. In either case, one extra mutation is sufficient to lead the dynamic to fall into a new absorbing set with strictly greater numbers of cooperators and punishers.

Following this path by induction, let the sets of cooperators and punishers increase step by step from the initial J_k and J'_k by taking an extra mutation at a time, going through absorbing sets $L_2, L_3, \ldots, L_{\kappa-1}$ and spreading cooperation and punishment to the entire circle population. Finally, $L_\kappa = \{z^C\}$ is reached. The key observation is that one extra mutation is required if and only if the dynamic falls into a new absorbing set $L \in \{L_2, L_3, \ldots, L_{\kappa-1}\}$ along the path. But $R(L) \ge 1$ for each $L \in \{L_2, L_3, \ldots, L_{\kappa-1}\}$ by definition. Therefore, the modified cost of such step-by-step increment path $(z_1, z_2, \ldots, z_k)$ from $z_1 = z^D$

to $z_k = z^C$ is $\tau^*(z_1, z_2, \dots, z_k) = \sum_{i=1}^{k-1} \tau(z_i, z_{i+1}) - \sum_{i=2}^{\kappa-1} R(L_i) = R(z^D)$. Since z^D is already the most distant state from z^C, we obtain $CR^*(z^C) \le R(z^D) = \lceil m\min\{\eta, \gamma^\dagger\}\rceil$.[28] □

Claim 2. *If $\eta + \gamma \ge 1 + \frac{1}{m}$, then $R(z^D) > n - m$.*

Proof of Claim 2. Let z^D be the initial state and $J = \{1, \dots, k\}$ be a consecutive set of individuals who make mistakes and take $x_i = y_i = 1$. Suppose $k \le n - m$, which means the mutations do not intersect with k's neighborhood M_k from the right-hand side. Consider the two cases and the construction of J, J', and J'' in Claim 1.

First, consider Case 1. To maintain the size of J, we need $\frac{|J\cap M_k|}{m} = \min\{\frac{k}{m}, \frac{m+1}{2m}\} > \eta$; otherwise, $J' \subset J$, and some individuals in J' would be indifferent or better off with alternative strategies, causing J to shrink over time and fall back to z^D. It is also necessary that $|J' \cap M_k| = m + 1 - \lceil m\eta\rceil > m\gamma$, which leads to $\lceil m\eta\rceil + m\gamma < m + 1$. However, $\eta + \gamma \ge 1 + \frac{1}{m}$ implies a contradiction: $\lceil m\eta\rceil + m\gamma > m\eta + m\gamma \ge m + 1$.

Second, consider Case 2. To leave $B(z^D)$, it is necessary that $\max\{\gamma, \hat{\gamma}(q^\dagger)\} < \frac{|J\cap M_k|}{m} = \min\{\frac{k}{m}, \frac{m+1}{2m}\}$, where $q^\dagger = \frac{|J''\cap M_k|}{m} = \frac{1}{m}(m + 1 - \lceil m\gamma\rceil)$, to ensure $J \subset J''$. Since $J \subset J''$, this implies $q^\dagger = \frac{|J''\cap M_k|}{m} \ge \frac{|J\cap M_k|}{m} = \min\{\frac{k}{m}, \frac{m+1}{2m}\}$. Therefore, $q^\dagger > \hat{\gamma}(q^\dagger)$. However, $\eta + \gamma \ge 1 + \frac{1}{m}$ implies $m\eta \ge m + 1 - m\gamma \ge m + 1 - \lceil m\gamma\rceil$, so that $q^\dagger = \frac{1}{m}(m + 1 - \lceil m\gamma\rceil) \le \eta$. Thus, a contradiction follows: $\hat{\gamma}(q^\dagger) \ge \hat{\gamma}(\eta) = \eta \ge q^\dagger$.

In either case, any $k \le n - m$ is insufficient to leave $B(z^D)$. Although we have only considered paths with a consecutive set of mutations, any path with non-consecutive mutations would require even more mistakes. Therefore, $R(z^D) > n - m$. □

Claim 3. *Suppose $m(1-\eta)$ and $m(1-\gamma)$ are not integers. Let $q^\ddagger = \frac{1}{m}(m - 1 - \lfloor m\gamma\rfloor)$ and $\gamma^\ddagger = \min\{\gamma, \hat{\gamma}(q^\ddagger)\}$. If $\eta + \gamma^\ddagger > 1$, then $CR^*(z^D) \le R(z^C) = \lceil m(1 - \max\{\eta, \gamma^\ddagger\})\rceil$.*

Proof of Claim 3. Let the initial state be z^C. The path with the minimum number of mistakes to escape $B(z^C)$ mirrors that of escaping $B(z^D)$. Given $\eta + \gamma > 1$, we have $(1-\eta) + (1-\gamma^\ddagger) < 1$, and one of $1-\eta$ and $1-\gamma^\ddagger$ is strictly less than $\frac{1}{2}$. First, let a set of individuals $J = \{1, 2, \dots, k\}$ defect and stop punishing due to mistakes.

In the case of $1 - \eta < \frac{1}{2} < \frac{m+1}{2m}$, let all individuals in $J' = \{i \in N | \frac{|J\cap M_i|}{m} > 1 - \eta\}$ update beliefs and best-respond by stopping punishment. For the sets J and J' not to shrink over time and fall back to z^C, it is sufficient and necessary that (i) $\frac{|J\cap M_k|}{m} > 1 - \eta$ and (ii) $\frac{|J'\cap M_k|}{m} = \frac{1}{m}(m + 1 - \lceil m(1-\eta)\rceil) > 1 - \gamma$. Condition (ii) is implied by $(1-\eta) + (1-\gamma^\ddagger) < 1$, since $m + 1 - \lceil m(1-\eta)\rceil > m + 1 - m(1-\eta) - 1 > m[1 - (1-\eta)] > m(1-\gamma^\ddagger) \ge m(1-\gamma)$. Condition (i) holds if $\frac{|J\cap M_k|}{m} = \min\{\frac{m+1}{2m}, \frac{k}{m}\} > 1 - \eta$. Therefore, the minimum k required to achieve this is $\lceil m(1-\eta)\rceil$.

In the case of $1 - \gamma^\ddagger < \frac{1}{2} < \frac{m+1}{2m}$, we have $1 - \gamma < \frac{m+1}{2m}$. First, let all individuals in $J = \{1, \dots, k\}$ make mistakes and switch to $x_i = y_i = 0$. Second, let all individuals in $J'' = \{i \in N | \frac{|J\cap M_i|}{m} > 1 - \gamma\}$ update their beliefs, best-respond, and defect. For J and J'' not to shrink over time, the sufficient and necessary conditions are that (i) $1 - p_k = \frac{|J\cap M_k|}{m} > 1 - \gamma$ to have $J \subset J''$ and (ii) $1 - p_k = \frac{|J\cap M_k|}{m} > 1 - \hat{\gamma}(q^\ddagger)$ to make individuals in J not punish, where

$$
\begin{aligned}
q^\ddagger &= 1 - \frac{|J''\cap M_k|}{m} \\
&= 1 - \frac{1}{m}(m + 1 - \lceil m(1-\gamma)\rceil) \\
&= \frac{1}{m}(\lceil m(1-\gamma)\rceil - 1) \\
&= \frac{1}{m}(m - 1 - \lfloor m\gamma\rfloor)
\end{aligned}
$$

is the fraction of cooperators in k's neighborhood M_k. Hence, it is required that $\frac{k}{m} > 1 - \gamma$ and $\frac{k}{m} > 1 - \hat{\gamma}(q^\ddagger)$. Thus, the minimum k is $\lceil m(1 - \min\{\gamma, \hat{\gamma}(q^\ddagger)\})\rceil$.

Combining both cases leads to $R(z^C) = \lceil m(1 - \max\{\eta, \gamma^\ddagger\})\rceil$, where $\gamma^\ddagger = \min\{\gamma, \hat{\gamma}(q^\ddagger)\}$.

Next, index the sets of defectors and non-punishers by k: J_k, J'_k, and J''_k. Denote $k' = |J'_k|$ and $k'' = |J''|$. Then, $|J_{k+1} \cap M_{k'+1}| = |J_k \cap M_{k'+1}| - 1 = |J_k \cap M_{k'}|$. Hence, $\frac{|J_k\cap M_{k'}|}{m} > 1 - \eta$ implies $\frac{|J_{k+1}\cap M_{k'+1}|}{m} > 1 - \eta$. Similarly, if $\frac{|J_k\cap M_{k'}|}{m} > \max\{1-\gamma, 1-\hat{\gamma}(q^\ddagger)\}$, then $\frac{|J_{k+1}\cap M_{k'+1}|}{m} > \max\{1-\gamma, 1-\hat{\gamma}(q^\ddagger)\}$. Hence, regardless of whether $1 - \eta < \frac{1}{2}$ or $1 - \gamma^\ddagger < \frac{1}{2}$, one extra mutation is sufficient to move the dynamic to a next absorbing set L with states including greater numbers of defectors and non-punishers. Through this process, we construct a step-by-step increment path $(z_1, z_2, \dots, z_k)$ from $z_1 = z^C$ to $z_k = z^D$ by induction and let $L_2, L_3, \dots, L_{\kappa-1}$ be the absorbing sets on the path besides $\{z^C\}$ and $\{z^D\}$. Then, $\tau^*(z_1, z_2, \dots, z_k) = \sum_{i=1}^{k-1} \tau(z_i, z_{i+1}) - \sum_{i=2}^{\kappa-1} R(L_i) = R(z^C)$. Since z^C is already the most distant state from z^D, we obtain $CR^*(z^D) \le R(z^C) = \lceil m(1 - \max\{\eta, \gamma^\ddagger\})\rceil$. □

Claim 4. *If $\eta + \gamma \le 1 - \frac{1}{m}$, then $R(z^C) > n - m$.*

Proof. Let z^C be the initial state and consider paths with minimum mistakes to escape $B(z^C)$. Suppose $k \le n - m$ and consider the paths in the two cases of Claim 3. In Case 1, for J and J' not to shrink over time, it is necessary that (i) $k > m(1-\eta)$ and

[28] It is $CR^*(z^C) \le R(z^D)$ rather than $CR^*(z^C) = R(z^D)$ because only a subset of paths from z^D to z^C has been considered.

(ii) $\frac{1}{m}(m+1-\lceil m(1-\eta)\rceil) > 1-\gamma$. The second condition leads to $m(1-\gamma)+\lceil m(1-\eta)\rceil < m+1$, while $\eta+\gamma \le 1-\frac{1}{m}$ implies $m(1-\gamma)+\lceil m(1-\eta)\rceil \ge m(1-\gamma)+m(1-\eta) \ge m+1$, resulting in a contradiction.

In Case 2, to leave $B(z^D)$, it is necessary that $\frac{|J\cap M_k|}{m} = \min\{\frac{k}{m}, \frac{m+1}{2m}\} > \max\{1-\gamma, 1-\hat{\gamma}(q^\ddagger)\}$, where $q^\ddagger = 1-\frac{|J''\cap M_k|}{m} = \frac{1}{m}(m-1-\lfloor m\gamma\rfloor)$, to ensure that $J\subset J''$ and that both J and J'' are sustainable. Note that $J\subset J''$ implies $\frac{|J''\cap M_k|}{m} \ge \frac{|J\cap M_k|}{m}$. Hence, $q^\ddagger \le 1-\max\{1-\gamma, 1-\hat{\gamma}(q^\ddagger)\} = \min\{\gamma, \hat{\gamma}(q^\ddagger)\} \le \hat{\gamma}(q^\ddagger)$. However, $\eta+\gamma\le 1-\frac{1}{m}$ implies $m(1-\gamma)+m(1-\eta)\ge m+1$, leading to $m(1-\eta)\ge m+1-m(1-\gamma)\ge m+1-\lceil m(1-\gamma)\rceil$. Hence, $1-q^\ddagger = \frac{1}{m}(m+1-\lceil m(1-\gamma)\rceil) \le 1-\eta$, so $q^\ddagger \ge \eta$. But then a contradiction follows: $\hat{\gamma}(q^\ddagger) \le \hat{\gamma}(\eta) = \eta \le q^\ddagger$.

In either case, any $k \le n-m$ is insufficient to leave $B(z^C)$. Although we have only considered paths with a consecutive set of mutations, any path with non-consecutive mutations would require even more mistakes. Therefore, $R(z^C) > n-m$. □

Proof of Theorem 3

Lemma 3. *Consider P^0 under mobility. Suppose $\eta < 1$. Then there are exactly four absorbing sets of P^0: $\mathcal{C}$, $\mathcal{D}$, $\mathcal{C}_1$, and $\mathcal{C}_2$.*

Proof of Lemma 3. For each state $z\in Z$ and $t\ge 1$, let $\Delta^t(z)\subset Z$ denote the set of states to which the dynamic P^0 under mobility can transition from z with positive probability in exactly t periods.

Step 1: $\mathcal{D}$ is an absorbing set of P^0 under mobility.

First, let the dynamic start from $z\in\mathcal{D}$. Consider an arbitrary individual $i\in N$. In the next period, for each $v\in V$, individual i has either beliefs $p'_{iv}=p_{iv}$ or $p'_{iv}=\frac{1}{|N_v|}\sum_{j\in N_v} y_j$. In either case, $p'_{iv}=0<\gamma$. Likewise, $q'_{iv}=0<\eta$. Hence, for any state z' in the next period, we have $x'_i=y'_i=0$ for each $i\in N$, so $z'\in\mathcal{D}$. Therefore, by induction, $\Delta^t(z)\subset\mathcal{D}$ for each $t\ge1$. Second, since the states in $\mathcal{D}$ differ only in the number of individuals in each community, and in all communities all individuals expect the defection norm and plan to take $x_i=y_i=0$, the expected payoffs for all communities are the same. Without mistakes, no one will change their behavior in any community. Therefore, when having the opportunity to choose a community, all individuals randomly pick one. It follows that for each z and z' from $\mathcal{D}$, there is a positive probability of transitioning from z to z' within one period by reassigning individuals accordingly through chance movement.

Step 2: If $\eta<1$, then $\mathcal{C}$ is an absorbing set of P^0 under mobility. The proof is analogous to that of Step 1 and is therefore omitted. The condition $\eta<1$ ensures that if $p_{iv}=q_{iv}=1$, then $p_{iv}=q_{iv}>\eta$, so that all individuals punish defectors in each $z\in\mathcal{C}$ and all future $z'\in\Delta^t(z)$.

Step 3: If $\eta<1$, then $\mathcal{C}_v$, for each $v\in V$, is an absorbing set of P^0 under mobility. Let P^0 start from a state $z\in\mathcal{C}_1$. Consider a state z' in the next period. For each $i\in N$, we have either $p'_{iv}=p_{iv}$ or $p'_{iv}=\frac{1}{|N_v|}\sum_{j\in N_v} y_j$ for each $v\in V$. In either case, $p'_{i1}=1$ while $p'_{i2}=0$. Similarly, $q'_{i1}=1$ and $q'_{i2}=0$. Hence, i takes $x'_i=y'_i=1$ if i lives in community 1, and $x'_i=y'_i=0$ if living in community 2. Thus, expecting a cooperation norm in community 1, the expected payoff for the community is $\frac{1}{2}(b+\bar{d})$, and expecting a defection norm in community 2, the expected payoff for community 2 is $\frac{1}{2}(\bar{b}+d)$. By the assumption of $b+\bar{d}>\bar{b}+d$, all mobile individuals will choose community 1. Thus, community 1 will have $n+m$ individuals in z'. Community 2 includes only the fixed residents; therefore, its size is $n-m$. Hence, $z'\in\mathcal{C}_1$. By induction, $\Delta^t(z)\subset\mathcal{C}_1$ for each $t\ge1$. Therefore, $\mathcal{C}_1$ is an absorbing set of P^0 under mobility. The same logic applies to $\mathcal{C}_2$.

To complete our proof, the final step demonstrates that no other absorbing sets exist.

Step 4: Starting from any state $z\in Z$, there is a positive probability of P^0 transitioning to $\mathcal{C}\cup\mathcal{D}\cup\mathcal{C}_1\cup\mathcal{C}_2$ within finite periods. Starting from z at period $t=0$, we require all individuals to update beliefs and have $p'_{iv}=\frac{1}{|N_v|}\sum_{j\in N_v} y_j$ and $q'_{iv}=\frac{1}{|N_v|}\sum_{j\in N_v} x_j$ in the next period, $t=1$. Then, for each $v\in V$, we have $q'_{iv}=q'_{jv}$ and $p'_{iv}=p'_{jv}$ for each $i,j\in N$. Thus, we can require all individuals to take the same best responses to q'_{iv} and p'_{iv} (in the case of indifference, we require them to cooperate and punish). It follows that all individuals in a community have the same beliefs and take the same actions at period $t=1$. Fix a $v\in V$. Then there are only four cases: (i) $p'_{iv}\ge\gamma$, $p'_{iv}\ge\hat{\gamma}(q'_{iv})$, and $x'_i=y'_i=1$ for each $i\in N'_v$; (ii) $p'_{iv}<\gamma$, $p'_{iv}<\hat{\gamma}(q'_{iv})$, and $x'_i=y'_i=0$ for each $i\in N'_v$; (iii) $p'_{iv}\ge\gamma$, $p'_{iv}<\hat{\gamma}(q'_{iv})$, $x'_i=1$ and $y'_i=0$ for each $i\in N'_v$; and (iv) $p'_{iv}<\gamma$, $p'_{iv}\ge\hat{\gamma}(q'_{iv})$, $x'_i=0$ and $y'_i=1$ for each $i\in N'_v$.

In cases (i) and (ii), all individuals update beliefs in the next period, so that in period $t=2$, either a cooperation equilibrium or a defection equilibrium is established in community v. In case (iii), consider a state z'' at period $t=2$ where all individuals update only their beliefs about the proportion of punishers, such that $p''_{iv}=\frac{1}{N'_v}\sum_{j\in N'_v} y'_i=0<\gamma$ while $q''_{iv}=q'_{iv}$ for each $i\in N$. Then $x''_i=y''_i=0$ for each $i\in N''_v$, leading to a defection equilibrium for community v in the next period, $t=3$. In case (iv), consider a state z'' at period $t=2$ where all individuals update only their beliefs about the proportion of cooperators, resulting in $q''_{iv}=\frac{1}{N'_v}\sum_{j\in N'_v} x'_i=0$ and $p''_{iv}=p'_{iv}<\gamma$ so that $x''_i=y''_i=0$ for each $i\in N''_v$ for each $i\in N$, reaching a defection equilibrium for community v in period $t=3$. Therefore, by applying this argument to each community $v\in V$, we can establish either a cooperation norm or a defection norm for each $v\in V$ within $t\le3$ periods, causing the dynamic to fall to a state $z\in\mathcal{C}\cup\mathcal{D}\cup\mathcal{C}_1\cup\mathcal{C}_2$. □

Let $\Omega=\{\mathcal{C},\mathcal{D},\mathcal{C}_1,\mathcal{C}_2\}$ be the collection of absorbing sets of P^0. In what follows, we present three lemmas that together imply Theorem 3. For two absorbing sets $L,L'\subset\Omega$, denote by $\tau(L,L')$ the minimum number of mistakes needed to transition from a state in L to a state in L'. Let $\alpha=\min\{\eta,\gamma^*\}$ and $\beta=\min\{1-\eta,1-\gamma^*\}$, where $\gamma^*=\max\{\gamma,\hat{\gamma}(1)\}$. Fig. A.2 presents the minimum numbers of mistakes required to transition from one absorbing set to another, which is summarized in Lemma 4.

Lemma 4. *Consider P^0 under mobility. Then*

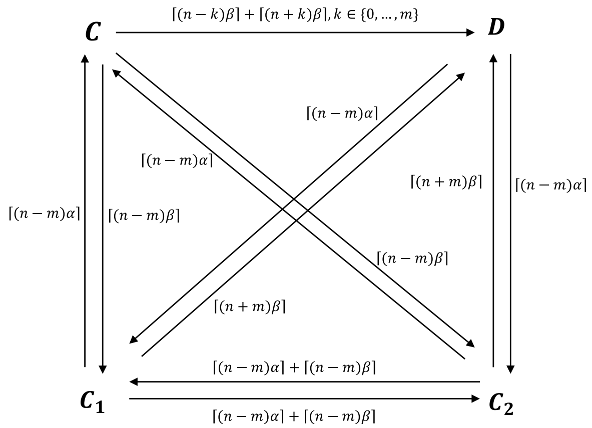

Fig. A.2. The costs (the minimum numbers of mistakes) for transitions between absorbing sets under mobility, where $\alpha = \min\{\eta, \gamma^*\}$ and $\beta = \min\{1-\eta, 1-\gamma^*\}$.

1. *$\tau(C, C_v) = \lceil (n-m)\beta \rceil$ while $\tau(C_v, C) = \lceil (n-m)\alpha \rceil$ for each $v \in V$;*
2. *$\tau(D, C_v) = \lceil (n-m)\alpha \rceil$ while $\tau(C_v, D) = \lceil (n+m)\beta \rceil$ for each $v \in V$;*
3. *$\tau(C_v, C_{v'}) = \tau(C_v, C) + \tau(C, C_{v'}) = \lceil (n-m)\alpha \rceil + \lceil (n-m)\beta \rceil$ for each $v, v' \in V$, $v \neq v'$;*
4. *$\tau(D, C) = \tau(D, C_1) + \tau(C_1, C) = 2\lceil (n-m)\alpha \rceil$ while $\tau(C, D) = \min\{\lceil (n-k)\beta \rceil + \lceil (n+k)\beta \rceil | k \in \{0,\dots,m\}\} \geq 2n\beta$.*

Proof of Lemma 4. To consider the transition from C to C_1, we start from a state in C and, in the first step, wait for all mobile individuals to move to community 1 by chance. This step requires no mistakes because all communities in a state in C enjoy the same expected utility, in which case each community is selected by each free individual with positive probability. Subsequently, community 2 contains only $n-m$ fixed residents. For community 2 to transition to a defection equilibrium in the next step, we consider two cases, $1-\eta \leq 1-\gamma^*$ and $1-\gamma^* < 1-\eta$, investigated in the proof of Theorem 1, and apply the two corresponding transition paths from z^C to z^D, respectively. Here, the sample size is $n-m$.

Specifically, when $1-\eta \leq 1-\gamma^*$, we allow $k = \lceil (n-m)(1-\eta) \rceil$ individuals to make mistakes and set $x_i = y_i = 0$ in community 2. Subsequently, all $i \in N_2$ update their beliefs to $q_i = p_i = \frac{n-m-k}{n-m} \leq \eta$. As a best-reply strategy, all $i \in N_2$ set $y_i = 0$. The individuals in community 1 maintain their original beliefs and remain in community 1 for the next several periods. Then, all $i \in N_2$ revise, but only revise, their belief about the proportion of punishers in the population, resulting in $p_i = 0 < \gamma$ and $q_i = \frac{n-m-k}{n-m} \leq \eta$. Consequently, all $i \in N_2$ set $x_i = y_i = 0$. When all $i \in N_2$ update their beliefs, community 2 transitions to a defection equilibrium in the subsequent period.

In the case of $1-\gamma^* < 1-\eta$, after $k = \lceil (n-m)(1-\gamma^*) \rceil$ individuals in community 2 take $x_i = y_i = 0$ by mistake, all $i \in N_2$ update beliefs to $q_i = p_i = \frac{n-m-k}{n-m} \leq \gamma^* = \max\{\gamma, \hat{\gamma}(1)\}$. If $\gamma^* = \gamma$, then $p_i \leq \gamma$ for each $i \in N_2$, and all of them would take $x_i = 0$ as a best response. In this case, in the next period, all individuals will revise beliefs to $q_i = 0$ and $p_i = \frac{n-m-k}{n-m} \leq \gamma$, taking $x_i = y_i = 0$ as a best response. If $\gamma < \gamma^* = \hat{\gamma}(1)$, then $q_i < 1$ and $p_i \leq \hat{\gamma}(1)$, so that we can require all $i \in N_2$ to take $y_i = 0$ as a best response. Then in the next period, all $i \in N_2$ update beliefs to $p_i = 0$ and $q_i = q_i = \frac{n-m-k}{n-m} \leq \hat{\gamma}(1) \leq \hat{\gamma}(1)$ and best-reply with $x_i = y_i = 0$. By requiring all individuals to update beliefs in the next period, community 2 reaches a defection equilibrium.

It follows that $\tau(C, C_1) = \lceil (n-m)\beta \rceil$, where $\beta = \min\{1-\eta, 1-\gamma^*\}$. Since C_1 is symmetric to C_2, $\tau(C, C_2) = \tau(C, C_1)$. By replacing the sample size with $n+m$, the same paths from z^C to z^D can be used to show that $\tau(C_v, D) = \lceil (n+m)\beta \rceil$. This equality holds because requiring $\lceil (n+m)\beta \rceil$ individuals in community v to make mistakes enables the community to transition to a defection equilibrium.

Similarly, to examine $C_v \to C$, we adapt the transition paths from z^D to z^C considered in the proof of Theorem 1. By replacing the sample size with $n-m$, we can show $\tau(C_v, C) = \lceil (n-m)\alpha \rceil$ for each $v \in V$, where $\alpha = \min\{\eta, \gamma^*\}$. Additionally, the same paths from z^D to z^C can be used to show $\tau(D, C_v) = \lceil (n-m)\alpha \rceil$: first, wait for all mobile individuals to move to community v'; then, make community v tip to a cooperation equilibrium by requiring $\lceil (n-m)\alpha \rceil$ individuals in community v to make mistakes.

To show $\tau(C_v, C_{v'}) = \tau(C_v, C) + \tau(C, C_{v'})$, it suffices to show that a path from C_1 to C_2 (or a path from C_2 to C_1) with a minimum number of mistakes must pass through C. To demonstrate this, observe that a path from C_1 to C_2 through C first allows community 2, with minimum size $n-m$, to tip to a cooperation equilibrium, thereby minimizing the number of mistakes needed for community 2 to change to a cooperation equilibrium. Second, once both communities achieve a cooperation equilibrium, we can wait for community 1 to reach the minimum size $n-m$ without mistakes and then require community 1 to change to a defection equilibrium with $\tau(C, C_2) = \lceil (n-m)\min\{1-\eta, 1-\gamma^*\} \rceil$ mistakes.

By the same argument, a path from D to C with the minimum number of mistakes is a path through C_1 or C_2. Therefore, $\tau(D, C) = \tau(D, C_1) + \tau(C_1, C)$. Similarly, a path from C to D with the minimum number of mistakes is either a path through C_1 or C_2, in which case $\tau(C, D) = \tau(C, C_1) + \tau(C_1, D)$, or a direct path to D where a sufficient number of individuals make mistakes and take $x_i = y_i = 0$ in both communities simultaneously. In the latter case, if the size of one community is $n+k$, then the size of the other community must be $n-k$, for $0 \leq k \leq m$. The total number of necessary mistakes is $\lceil (n-k)\beta \rceil + \lceil (n+k)\beta \rceil$. Consequently, $\tau(C, D) = \min\{\lceil (n-k)\beta \rceil + \lceil (n+k)\beta \rceil | k \in \{0,\dots,m\}\} \geq 2n\beta$. ☐

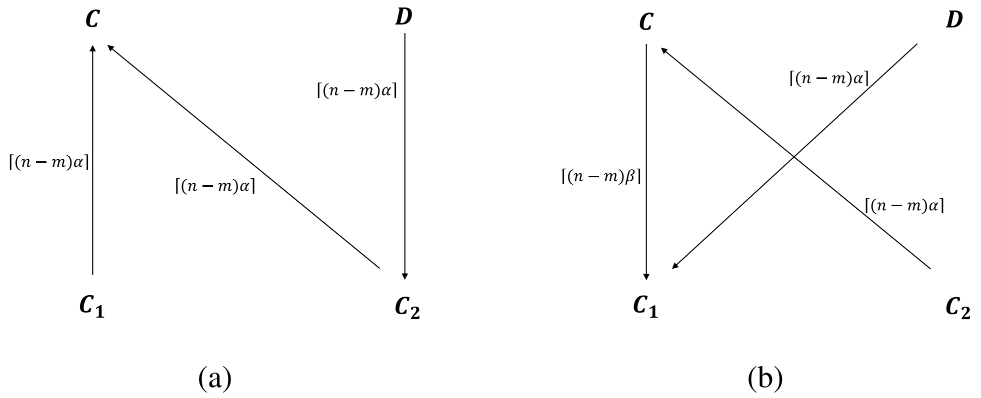

Fig. A.3. Assuming $\lceil (n-m)\beta \rceil < \lceil (n-m)\alpha \rceil < \lceil (n+m)\beta \rceil$, (a) is the stochastic tree that minimizes costs over all stochastic trees rooted at C and (b) is one minimizing costs over all stochastic trees rooted at C_1.

Lemma 4 identifies the minimum number of mistakes required for transitions between any two absorbing sets in Ω. By applying either Young's (1993) theorem or Ellison's (2000) theorem, this suffices to determine the stochastically stable sets. First, I establish the following observation:

Lemma 5. *$\eta + \gamma^* \geq 1$ if and only if $\alpha \geq \beta$.*

Proof of Lemma 5. Notice that $\eta + \gamma^* = \min\{\eta, \gamma^*\} + \max\{\eta, \gamma^*\}$. Hence, $\eta + \gamma^* \geq 1$ if and only if

$$\begin{aligned}
\alpha - \beta &= \min\{\eta, \gamma^*\} - \min\{1-\eta, 1-\gamma^*\} \\
&= \min\{\eta, \gamma^*\} - (1 - \max\{\eta, \gamma^*\}) \\
&= \min\{\eta, \gamma^*\} + \max\{\eta, \gamma^*\} - 1 \\
&\geq 0,
\end{aligned}$$

completing the proof of Lemma 5. □

Lemma 6. *Consider P^0 under mobility. Suppose $\lceil (n-m)\beta \rceil \neq \lceil (n-m)\alpha \rceil$. If $\eta + \gamma^* < 1$, then the long-run stable equilibria are the states in $\mathcal{C}$.*

Proof of Lemma 6. From Lemma 4, we know that the radii of $\mathcal{C}$ and C_1 are $R(\mathcal{C}) = \lceil (n-m)\beta \rceil$ and $R(C_1) = \lceil \min\{(n-m)\alpha, (n+m)\beta\} \rceil$. Given $\eta + \gamma^* < 1$, we have $\alpha < \beta$. Hence, $R(C_1) = \lceil (n-m)\alpha \rceil$. It follows that the modified co-radius of $\mathcal{C}$ is $CR^*(\mathcal{C}) = \tau(\mathcal{D}, C_1) + \tau(C_1, \mathcal{C}) - R(C_1) = \lceil (n-m)\alpha \rceil$. According to Ellison's (2000) theorem, the states in $\mathcal{C}$ are uniquely stochastically stable if $R(\mathcal{C}) > CR^*(\mathcal{C})$, which holds if and only if $\lceil (n-m)\beta \rceil > \lceil (n-m)\alpha \rceil$, which is implied by $\alpha < \beta$ and $\lceil (n-m)\beta \rceil \neq \lceil (n-m)\alpha \rceil$. □

Lemma 7. *Consider P^0 under mobility. Suppose $\lceil (n-m)\alpha \rceil \neq \lceil (n+m)\beta \rceil$. If $\eta + \gamma^* > 1 + \frac{m}{n}(1 - |\eta - \gamma^*|)$ such that $\lceil (n-m)\alpha \rceil > \lceil (n+m)\beta \rceil$, then the long-run stable equilibria are the states in $\mathcal{D}$.*

Proof of Lemma 7. According to Lemma 4, the radius of $\mathcal{D}$ is $R(\mathcal{D}) = \lceil (n-m)\alpha \rceil$, and the radius of C_1 is $R(C_1) = \lceil \min\{(n-m)\alpha, (n+m)\beta\} \rceil = \lceil (n+m)\beta \rceil$ given $\lceil (n-m)\alpha \rceil > \lceil (n+m)\beta \rceil$. Therefore, $CR^*(\mathcal{D}) \leq \tau(\mathcal{C}, C_1) + \tau(C_1, \mathcal{D}) - R(C_1) = \lceil (n-m)\beta \rceil < \lceil (n-m)\alpha \rceil$, which, according to Ellison's (2000) theorem, implies that the states in $\mathcal{D}$ are uniquely stochastically stable. □

Lemma 8. *Consider P^0 under mobility. Suppose $\lceil (n-m)\alpha \rceil \neq \lceil (n-m)\beta \rceil$ and $\lceil (n-m)\alpha \rceil \neq \lceil (n+m)\beta \rceil$. If $1 < \eta + \gamma < 1 + \frac{m}{n}(1 - |\eta - \gamma|)$ so that $\lceil (n-m)\beta \rceil < \lceil (n-m)\alpha \rceil < \lceil (n+m)\beta \rceil$, then the long-run stable equilibria are the states in $C_1 \cup C_2$.*

Proof of Lemma 8. The proof for this case relies on finding the stochastic tree with minimum cost over all stochastic trees defined on Ω and then applying Young (1993)'s theorem. Finding the stochastic tree with minimum cost is equivalent to solving the Minimum Spanning Arborescence problem in graph theory. By applying the established Chu-Liu-Edmonds' algorithm (Chu, 1965; Edmonds, 1967), we narrow down to three candidate trees: one rooted at $\mathcal{C}$ and two rooted at C_1 and C_2 respectively. The tree rooted at $\mathcal{C}$ and the one rooted at C_1 are displayed in Fig. A.3. The former has a cost of $3\lceil (n-m)\alpha \rceil$ while the latter has a cost of $2\lceil (n-m)\alpha \rceil + \lceil (n-m)\beta \rceil$. Given $\eta + \gamma > 1$ and $\lceil (n-m)\alpha \rceil \neq \lceil (n-m)\beta \rceil$ so that $\lceil (n-m)\alpha \rceil > \lceil (n-m)\beta \rceil$, we conclude that the stochastic tree rooted at C_1 has minimum cost. We can construct a stochastic tree rooted at C_2 symmetric to that of C_1 with the same cost. It follows that the long-run stable equilibria are the states in $C_1 \cup C_2$. □

Online Appendix

This Online Appendix provides supplementary details and extensions to Xueheng Li's paper "Indignation and the Evolution of Cooperation Norms" accepted by Games and Economic Behavior. It includes three sections: first, a general model of indignation in dynamic games; second, a dynamic framework for general psychological games; and third, local interactions on star networks.

I. A General Model of Indignation

The model is a *population game* among a large number of individuals, $N = \{1, 2, \dots, n\}$, who are randomly grouped to play a *base game*. The base game is a finite extensive-form game of perfect information, represented by a game tree with material payoff numbers at terminal nodes. The base game represents instances of daily interactions among unrelated individuals in society, such as queuing at a bus stop. The set H includes all decision nodes in the game tree. Each $i \in N$ is assigned a unique decision node $h \in H$, which is interpreted as i's *player role*.[29] The material payoffs for an individual in player role h when the game ends at a terminal node h^* of the game tree are given by $\pi_h(h^*)$.

Indignation is belief-dependent. Each individual has a belief about the *normal behavior* of a random individual in society. Let A_h denote the set of available actions at decision node h, for each $h \in H$. The belief, denoted by β_i for individual $i \in N$, is a conditional probability function, where $\beta_i(a|h)$ denotes i's belief about the probability of a randomly selected individual from the population taking action $a \in A_h$ at $h \in H$.[30]

Suppose that i acts at decision node h, i.e., i is assigned player role h. Then, for each $\ell \in H$, conditional on the game reaching ℓ and based on i's belief β_i, individual i's expected payoffs at the end of the game are denoted by $\mathbb{E}[\pi_h|\ell, \beta_i]$. In particular, if the game reaches i's decision node h and i chooses action $a_i \in A_h$, then her expected material payoffs are $\mathbb{E}[\pi_h|(h, a_i), \beta_i]$.[31]

Initially, individual i's expectation about her payoffs is $\mathbb{E}[\pi_h|\varnothing, \beta_i]$, where $\varnothing$ denotes the root node of the game tree. Based on i's initial belief, she might not need to take any action if she initially believes that the decision node h would not be reached. However, if a co-player at a predecessor node $\ell \in H$, referred to as co-player ℓ, takes an action different from i's initial expectation, then i may find herself reaching the unexpected node h, and the deviation leads to a loss: $\mathbb{E}[\pi_h|h, \beta_i] < \mathbb{E}[\pi_h|\varnothing, \beta_i]$. In such a situation, i becomes angry. When evaluating how much co-player ℓ is responsible for i's loss, individual i considers the counterfactual action that ℓ would have taken according to i's initial belief rather than the action that ℓ actually took. Formally, suppose that co-player ℓ takes action $a_\ell \in A_\ell$. Let $[r]^+ = \max\{0, r\}$ for any number r. Then i's *indignation at h toward co-player ℓ* is defined by

$$\rho_{h\ell}(\beta_i) = \Big[\mathbb{E}[\pi_h \mid \ell, \beta_i] - \mathbb{E}[\pi_h \mid (\ell, a_\ell), \beta_i]\Big]^+. \tag{A.1}$$

Individuals' decisions are dictated by their *decision utilities*. Individual i's decision utility for choosing $a_i \in A_h$ at h is

$$u_i(a_i|h, \beta_i) = \mathbb{E}[\pi_h|(h, a_i), \beta_i] - \lambda_i \sum_{\ell \in H \setminus \{h\}} \rho_{h\ell}(\beta_i)\mathbb{E}[\pi_\ell|(h, a_i), \beta_i],$$

where $\lambda_i \geq 0$ measures i's primitive psychological sensitivity to indignation. The multiplication between indignation $\rho_{h\ell}(\beta_i)$ and player ℓ's payoffs implies that if the indignation against ℓ is positive and sufficiently strong, i is willing to reduce ℓ's payoffs at a personal cost.

An individual's strategy s_i in the population game is an action plan that specifies an action for each decision node in the base game (since player roles are randomly determined), where $s_i(h) \in A_h$ denotes i's action at $h \in H$.[32] Let $s = (s_i)_{i \in N}$ collect all individuals' strategies and $\beta = (\beta_i)_{i \in N}$ collect all individuals' beliefs. The pair (s, β) is called an *assessment*, or a *(population) state*, as it is the state variable of the dynamics examined later. A state (s, β) is said to be *consistent* if for each $h \in H$ and $a \in A_h$, every $i \in N$ holds a correct belief about the probability of action $a \in A_h$ chosen at h: $\beta_i(a|h) = \sum_{j \in N} \mathbf{1}\{s_j(h) = a\}/n$. An individual's strategy s_i is *sequentially rational relative to belief* β_i if action $s_i(h) \in A_h$ maximizes $u_i(h, s_i(h), \beta_i)$ for each $h \in H$. I use the term *indignation equilibrium* to refer to sequential equilibrium (Kreps and Wilson, 1982) in games with indignation decision utilities.

Definition 2. A state (s, β) is an *indignation equilibrium* if it is consistent and s_i is sequentially rational relative to β_i for each $i \in N$.

Remark. We have $\rho_{h\ell}(\beta_i) = 0$ in two situations. First, indignation is an ex-post phenomenon. In particular, $\rho_{\varnothing\ell}(\beta_i) = 0$ at the beginning of the game regardless of beliefs. Hence, a first-mover always maximizes expected material payoffs. Second, we have $\rho_{h\ell}(\beta_i) = 0$ if either (i) a co-player ℓ takes the action that individual i at h expects ℓ to take with probability one, or (ii) ℓ's deviation from i's initial belief increases rather than reduces i's expected payoff.

II. A General Dynamic Framework for Psychological Games

Consider a population $N = \{1, \dots, n\}$ of individuals who are recurrently matched to play a base game G. Let S denote the set of all possible strategy profiles $s = (s_1, \dots, s_n)$, and B denote the set of all possible belief profiles $\beta = (\beta_1, \dots, \beta_n)$, where β_i is the belief space for player i. The beliefs $\beta_i \in B_i$ may concern the strategies of immediate interaction partners, the strategy profile of the population, and higher-order beliefs about others' beliefs. Unlike standard games where payoffs depend only on strategies, each individual i has a belief-dependent utility function $u_i(s_i, s_{-i}, \beta_i)$. Let Z be the collection of all population states $z = (s, \beta)$.

[29] Having players assigned to different roles is a common feature in evolutionary games and stochastic dynamics, such as in Young (1993). Without much loss of generality, I assume n being a multiple of $|H|$.

[30] I consider the situation where n is much larger than $|H|$, so that whether excluding i's own action plan in forming beliefs about the population's aggregate properties has a negligible effect.

[31] Nodes in a game tree map one-to-one to the paths leading from the root node to the nodes under consideration. Thus, $h' = (h, a_i)$ denotes the successor node of h appended by action a_i.

[32] For the purpose of this paper, I focus on pure strategies and pure strategy equilibrium, and leave for future research mixed strategies whose interpretations are known to be subtle for psychological games; see discussions by Rabin (1993), Dufwenberg and Kirchsteiger (2004).

The dynamics unfold in discrete time periods. Let $BR_i(\beta_i^t) \subset S_i$ denote the set of best responses conditional on beliefs β_i^t; if the base game is a dynamic psychological game, $BR_i(\beta_i^t)$ denotes the set of sequentially rational pure strategies conditional on the belief hierarchy β_i^t (see Battigalli and Dufwenberg (2009) for the precise definition and proper treatments). In each period t, individuals take a best response to their beliefs: $s_i^t \in BR_i(\beta_i^t)$. Each period, each individual independently updates their beliefs with a positive probability according to a certain belief updating protocol. This protocol specifies how each individual draws observations from the strategy profile s^{t-1} in the last period and how these observations combine with their previous beliefs β_i^{t-1} to produce beliefs β_i^t in the current period. The belief updating protocol is summarized by the *set of admissible beliefs,* $\Theta_i(s, \beta_i) \subset B_i$, conditional on the previous state (s, β_i), meaning that beliefs $\beta_i' \in B_i$ can be formed with positive probability in the next period if and only if $\beta_i' \in \Theta_i(s, \beta_i)$.

In the dynamic without mistakes, individuals play best responses to their beliefs. Let $\varepsilon \in [0, 1)$ be the probability of making mistakes, and let $P^\varepsilon(z, z') \in [0, 1]$ be the probability of transitioning from state $z = (s, \beta)$ to $z' = (s', \beta')$ within a period. Then the best-response dynamic $P^0(z, z')$ satisfies:

Assumption 1. $P^0(z, z') > 0$ if and only if $\beta_i' \in \Theta_i(s, \beta_i)$ and $s_i' \in BR_i(\beta_i')$ for each $i \in N$.

Assume that mistakes occur in the form of actions but not in the belief formation process; all potential biases in the belief formation process have been summarized by $\Theta_i(s, \beta_i)$. Thus, even when $\varepsilon > 0$, direct transitions from $z = (s, \beta)$ to $z' = (s', \beta')$ remain impossible if $\beta_i' \notin \Theta_i(s, \beta_i)$. Let $d(z, z')$ denote the *resistance* of transitioning from z to z' within one period, defined as

$$d(z, z') = \begin{cases} \#\{i \in N | s_i' \notin BR_i(\beta_i')\} & \text{if } \beta_i' \in \Theta_i(s, \beta_i) \forall i \in N \\ \infty & \text{otherwise.} \end{cases}$$

The dynamic P^ε under $\varepsilon > 0$ is the perturbed version of the best-response dynamic Q^0.

Assumption 2. $P^\varepsilon(z, z') = P^0(z, z') + o(\varepsilon^{d(z,z')})$ for each $z, z' \in Z$.[33]

Thus, $\lim_{\varepsilon \to 0} P^\varepsilon(z, z') = P^0(z, z')$, and $P^\varepsilon(z, z') > 0$ if $\varepsilon > 0$ and $d(z, z') < \infty$, even when $P^0(z, z') = 0$. Nevertheless, even with mistakes, transitioning from a state z to state z' occurs with positive probability only if z' contains only admissible beliefs conditional on z, i.e., $\beta_i' \in \Theta_i(s, \beta_i)$ for each $i \in N$.

Let μ^ε be the unique stationary distribution of the perturbed process (if it exists). A state z is *long-run stable* if $\lim_{\varepsilon \to 0} \mu^\varepsilon(z) > 0$. Based on Young (1993) and Ellison (2000), long-run stable states depend on the minimum costs of transitions between states, as measured by the minimum number of necessary mistakes for the transitions to occur. Define by $\tau(z^1, \ldots, z^K)$ the cost of a transition path $(z^1, \ldots, z^K)$:

$$\tau(z^1, \ldots, z^K) = \sum_{k=1}^{K-1} d(z^k, z^{k+1}).$$

Let $\tau^*(z, z')$ denote the minimum cost of transitioning from state z to state z' within finite periods:

$$\tau^*(z, z') = \min_{(z^1, \ldots, z^K)} \tau(z^1, \ldots, z^K),$$

where the minimum is taken over all transition paths $(z^1, \ldots, z^K)$ from $z^1 = z$ to $z^K = z'$.

Note that if there exist two consecutive states $z^k = (s, \beta)$ and $z^{k+1} = (s^{k+1}, \beta^{k+1})$ such that $\beta_i^{k+1} \notin \Theta_i(s^k, \beta_i^k)$ on a path, then the cost of the path is infinitely high because $d(z^k, z^{k+1}) = \infty$. Thus, to minimize transition costs, it is necessary and sufficient to examine only the *admissible transition paths* - those containing only admissible beliefs for each state along the path except for the initial state. To be precise, the minimum transition cost $\tau^*(z, z')$ from state z to state z' is determined by solving

$$\min_{(z_1, \ldots, z_K)} \sum_{k=2}^{K} \#\{i \in N | s_i^k \notin BR_i(\beta_i^k)\}$$

subject to $z_1 = z$, $z_K = z'$, and

$$\beta_i^{k+1} \in \Theta_i(s^k, \beta_i^k) \text{ for each } k = 1, \ldots, K-1. \tag{A.2}$$

In other words, the task is to search for an admissible path that minimizes the number of non-best responses.

Remarks. In the models of Young (1993, 1998), Kandori et al. (1993), Ellison (1993), and Blume (1993), even when beliefs are explicitly written down, beliefs in the current period are assumed dependent on the strategy profile s in the previous period(s) but *not* directly on past beliefs β_i. Past beliefs will affect future beliefs and play only through their influence on concurrent best-responses. The distinguishing feature of my model is that beliefs that can be formed, $\Theta_i(s, \beta_i)$, may depend on both the past strategy profile s and the past beliefs β_i. This creates a direct path-dependence in beliefs, beyond their impact on concurrent best-responses. In the following, I briefly discuss (i) the importance of allowing for this direct path-dependence in beliefs in analyzing any non-trivial psychological games and (ii) its effects on transition costs between equilibria and the characterization of long-run stable states.

First, any non-trivial psychological games involve higher-order beliefs, raising questions about their updating rules over time. While other updating protocols are conceivable, one natural protocol is a bottom-up process: first-order beliefs are updated traditionally through sampling past strategy profiles and myopically computing the strategy distribution from the samples. However,

[33] I write $f(x) = g(x) + o(h(x))$ if there are constants $c_1, c_2 > 0$ such that $c_1 h(x) \leq f(x) - g(x) \leq c_2 h(x)$.

second-order beliefs are derived from sampling *first-order beliefs* in the previous period, third-order beliefs are derived from sampling previous *second-order beliefs*, and so on. The idea is that people's higher-order beliefs do not change instantaneously and simultaneously with their actions. Instead, beliefs at each order are communicated and revealed gradually, sequentially shifting higher-order beliefs. To model this bottom-up belief updating protocol, it is essential to allow current beliefs to directly depend on past beliefs.

Second, path-dependence in beliefs can modify transition costs between equilibria. In traditional models, the cost of a transition is determined solely by the number of required strategy deviations. However, if current beliefs can directly depend on past beliefs, then even with an identical number of initial strategy deviations, the transition cost may vary depending on the specific belief updating protocol assumed. This occurs because different protocols may lead to different sets of admissible beliefs, even given the same previous strategy profile, thereby altering the set of admissible transition paths and modifying the constraint (A.2) in searching for a minimum cost path. For example, in the analysis of cooperation and punishment under indignation over time, the transition cost from z^D to z^C depends on whether individuals may update only one component in their beliefs, (q_i, p_i). To illustrate, consider $m = n$. First, suppose that as in the traditional approach, future beliefs depend only on previous strategy profiles, implying that all belief components must update simultaneously. Let k be the initial number of strategy deviations from (defect, not punish) to (cooperate, punish). Then the beliefs (q_i', p_i') in the next period are $q_i' = p_i' = \frac{k}{n}$. Note that $q_i' = p_i' \geq \min\{\gamma, \eta\}$ is not sufficient for reaching the basin of attraction of z^C. There is a recurrent class containing a state in which everyone cooperates and does not punish and a state in which everyone defects while they all punish defectors. Either $\gamma \leq q_i' = p_i' < \eta$ or $\eta \leq q_i' = p_i' < \gamma$ would lead the dynamic to fall into this recurrent class in the next period. Thus, the minimum number of mistakes to reach z^C must be strictly greater than $\lceil n \min\{\gamma, \eta\} \rceil$. In contrast, if the dynamic follows the protocol considered in my main analysis - there may be path-dependence in beliefs, and with positive probability individuals update either q_i or p_i but not both in a period - then according to Lemmas 1 and 2, we have $\tau^*(z^D, z^C) = \lceil n \min\{\gamma, \eta\} \rceil$.

III. Star networks

Consider local interactions on a star network. Individual $i^* \in N$ is the star at the center of the network, whose accessible sample is $M_{i^*} = N \backslash \{i^*\}$ in belief updating. The remaining individuals $i \neq i^*$ are peripheral, each with an accessible sample of $M_i = \{i^*\}$. Let z^{SC} denote the state where the *s*tar *c*ooperates but does not punish while all peripheral individuals defect but punish, and they have correct beliefs. Let z^{SD} denote the opposite state in which the *s*tar *d*efects but punishes while peripheral individuals cooperate but do not punish, with correct beliefs.

Proposition 2. *Consider local interactions on a star network. Suppose* $0 < \eta < 1$ *and* $\hat{\gamma}(1) \neq 0$*. Then both* z^C *and* z^D *are long-run stable. If* $\hat{\gamma}(1) < 0$*, then* z^C, z^D, z^{SC} *and* z^{SD} *are all long-run stable.*

Proof. Suppose $\hat{\gamma}(1) > 0$. Then the dynamic without mistakes has only two recurrent classes: $\{z^C\}$ and $\{z^D\}$. To show that both z^C and z^D are long-run stable, we show that it takes only one mistake to transition from z^D to z^C, and likewise one mistake to transition back from z^C to z^D. Starting from z^D, let the star i^* change to $x_{i^*} = y_{i^*} = 1$ by mistake. In the next period, let all $i \neq i^*$ update beliefs to $p_i = q_i = 1$, so that $p_i > \gamma$ and $q_i > \hat{\gamma}(1)$, leading to $x_i = y_i = 1$. In the next period, peripheral individuals do not update beliefs, while the star updates beliefs to $p_{i^*} = q_{i^*} = 1$ and best-responds with $x_{i^*} = y_{i^*} = 1$, thus reaching z^C. Therefore, $R(z^D) = 1$, and the minimum transition cost from z^D to z^C is one. Following the same pattern of transition (by first requiring the star to cooperate and punish), we can show that the minimum transition cost from z^C to z^D is also one. According to Ellison (2000, p. 38, Theorem 3), these results imply both $\mu^*(z^C) > 0$ and $\mu^*(z^D) > 0$, where $\mu^* = \lim_{\epsilon \to 0} \mu^\epsilon$.

For the case of $\hat{\gamma}(1) < 0$, it is necessary to consider two additional recurrent classes, $\{z^{SC}\}$ and $\{z^{SD}\}$, where z^{SC} denotes the state in which the star cooperates but does not punish while all peripheral individuals defect but punish, and z^{SD} denotes the opposite state in which the star defects but punishes while peripheral individuals cooperate but do not punish. The condition $\hat{\gamma}(1)$ ensures that these two states are each absorbing, thus $R(z^{SC}) \geq 1$ and $R(z^{SD}) \geq 1$. By constructing a similar transition path as above (requiring the star to initiate a change), we note that the minimum number of mistakes needed to transition from each of the four absorbing states to another is exactly one. Thus, z^C, z^D, z^{SC}, and z^{SD} are all long-run stable. □

Thus, unlike the circle network in the main text, different norms can be simultaneously stable in star networks. This fundamental difference stems from the stark asymmetry in the star network structure, where the central player can observe everyone while peripheral players can only observe the center. Therefore, while the circle network provides a clean benchmark for comparison with previous literature, the star network results show that the network structure can fundamentally affect the long-run dynamics of cooperation under indignation, suggesting that a comprehensive investigation of network structures would be a valuable direction for future research.

References

Abramson, G., Kuperman, M., 2001. Social games in a social network. Phys. Rev. E 63 (3), 030901. https://doi.org/10.1103/PhysRevE.63.030901

Adriani, F., Sonderegger, S., 2009. Why do parents socialize their children to behave pro-socially? an information-based theory. J. Public Econ. 93 (11–12), 1119–1124. https://doi.org/10.1016/j.jpubeco.2009.08.001

Akerlof, G.A., 1982. Labor contracts as partial gift exchange. Q. J. Econ. 97 (4), 543–569. https://doi.org/10.2307/1885099

Akerlof, R., 2016. Anger and enforcement. J. Econ. Behav. Organ. 126, 110–124. https://doi.org/10.1016/j.jebo.2015.10.022

Alós-Ferrer, C., Weidenholzer, S., 2006. Imitation, local interactions, and efficiency. Econ. Lett. 93 (2), 163–168. https://doi.org/10.1016/j.econlet.2006.04.006

Alós-Ferrer, C., Weidenholzer, S., 2008. Contagion and efficiency. J. Econ. Theory 143 (1), 251–274. https://doi.org/10.1016/j.jet.2007.12.003

Anwar, A., 2002. On the co-existence of conventions. J. Econ. Theory 107 (1), 145–155. https://doi.org/10.1006/jeth.2001.2872

Battigalli, P., Dufwenberg, M., 2007. Guilt in games. Am. Econ. Rev. 97 (2), 170–176. https://doi.org/10.1257/aer.97.2.170

Battigalli, P., Dufwenberg, M., 2009. Dynamic psychological games. J. Econ. Theory 144 (April), 1–35. https://doi.org/http://dx.doi.org/10.1016/j.jet.2008.01.004

Battigalli, P., Dufwenberg, M., Smith, A., 2019. Frustration, aggression, and anger in leader-follower games. Games Econ. Behav. 117, 15–39.
Bergstrom, T.C., Stark, O., 1993. How altruism can prevail in an evolutionary environment. Am. Econ. Rev. 83 (2), 149–155.
Bhaskar, V., Vega-Redondo, F., 2004. Migration and the evolution of conventions. J. Econ. Behav. Organ. 55 (3), 397–418. https://doi.org/10.1016/j.jebo.2002.03.001
Bicchieri, C., 2006. The Grammar of Society: The Nature and Origins of Social Norms. Cambridge University Press. https://doi.org/10.1017/CBO9780511616037
Bisin, A., Verdier, T., 2000. "Beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits. Q. J. Econ. 115 (3), 955–988. https://doi.org/10.1162/003355300554953
Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. J. Econ. Theory 97 (2), 298–319. https://doi.org/10.1006/jeth.2000.2678
Bisin, A., Verdier, T., 2011. The economics of cultural transmission and socialization. In: Handbook of Social Economics. Elsevier B.V. Vol. 1, pp. 339–416. https://doi.org/10.1016/B978-0-444-53187-2.00009-7
Blume, A., Temzelides, T., 2003. On the geography of conventions. Econ. Theory 22 (4), 863–873. https://doi.org/10.1007/s00199-002-0350-3
Blume, L.E., 1993. The statistical mechanics of strategic interaction. Games Econ. Behav. 5 (3), 387–424. https://doi.org/10.1006/game.1993.1023
Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. Vol. 175. University of Chicago Press. https://doi.org/10.1097/00005053-198702000-00018
Camerer, C., Thaler, R.H., 1995. Anomalies: ultimatums, dictators and manners. J. Econ. Perspect. 9 (2), 209–219. https://doi.org/10.1257/jep.9.2.209
Carpenter, J.P., Matthews, P.H., 2012. Norm enforcement: anger, indignation, or reciprocity? J. Eur. Econ. Assoc. 10 (3), 555–572. https://doi.org/10.1111/j.1542-4774.2011.01059.x
Chu, Y.-J., 1965. On the shortest arborescence of a directed graph. Sci. Sin. 14, 1396–1400.
Cooper, D.J., Dutcher, E.G., 2011. The dynamics of responder behavior in ultimatum games: a meta-study. Exp. Econ. 14 (4), 519–546. https://doi.org/10.1007/s10683-011-9280-x
Cui, Z., Weidenholzer, S., 2021. Lock-in through passive connections. J. Econ. Theory 192, 105187. https://doi.org/10.1016/j.jet.2021.105187
Dieckmann, T., 1999. The evolution of conventions with mobile players. J. Econ. Behav. Organ. 38 (1), 93–111. https://doi.org/10.1016/S0167-2681(98)00125-5
Dixit, A.K., 2007. Lawlessness and Economics: Alternative Modes of Governance. Princeton University Press, Princeton, NJ. 1st edition.
Dufwenberg, M., Kirchsteiger, G., 2004. A theory of sequential reciprocity. Games Econ. Behav. 47 (2), 268–298. https://doi.org/10.1016/j.geb.2003.06.003
Durieu, J., Solal, P., 2003. Adaptive play with spatial sampling. Games Econ. Behav. 43 (2), 189–195. https://doi.org/10.1016/S0899-8256(03)00012-5
Edmonds, J., 1967. Optimum branchings. J. Res. Natl. Bureau Stand. B 71 (4), 233–240.
Ellison, G., 1993. Learning, local interaction, and coordination. Econometrica 61 (5), 1047–1071. https://doi.org/10.2307/2951493
Ellison, G., 2000. Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. Rev. Econ. Stud. 67 (1), 17–45. https://doi.org/10.1111/1467-937X.00119
Elster, J., 1989. Social norms and economic theory. J. Econ. Perspect. 3 (4), 99–117. https://doi.org/10.1257/jep.3.4.99
Elster, J., 1998. Emotions and economic theory. J. Econ. Lit. 36 (1), 47–74.
Ely, J.C., 2002. Local conventions. B.E. J. Theor. Econ. 2 (1), 1. https://doi.org/10.2202/1534-5963.1044
Engelmann, D., Strobel, M., 2004. Inequality aversion, efficiency, and maximin preferences in simple distribution experiments. Am. Econ. Rev. 94 (4), 857–869. https://doi.org/10.1257/0002828042002741
Eshel, I., Samuelson, L., Shaked, A., 1998. Altruists, egoists, and hooligans in a local interaction model. Am. Econ. Rev. 88 (1), 157–179. https://doi.org/10.2307/116823
Fehr, E., Fischbacher, U., Gächter, S., 2002. Strong reciprocity, human cooperation, and the enforcement of social norms. Hum. Nat. 13 (1), 1–25. https://doi.org/10.1007/s12110-002-1012-7
Foster, D., Young, H.P., 1990. Stochastic evolutionary game dynamics. Theor. Popul. Biol. 38 (2), 219–232. https://doi.org/10.1016/0040-5809(90)90011-J
Geanakopolos, J.P., Pearce, D., Stacchetti, E., 1989. Psychological game and sequential rationality. Games Econ. Behav. 1 (1), 60–79. https://doi.org/http://dx.doi.org/10.1016/0899-8256(89)90005-5
Gintis, H., 2000. Strong reciprocity and human sociality. J. Theor. Biol. 206 (2), 169–179. https://doi.org/10.1006/jtbi.2000.2111
Grujic, J., Fosco, C., Araujo, L., Cuesta, J.A., Sanchez, A., 2010. Social experiments in the mesoscale: humans playing a spatial prisoner's dilemma. PLoS ONE 5 (11), e13749. https://doi.org/10.1371/journal.pone.0013749
Güth, W., Schmittberger, R., Schwarze, B., 1982. An experimental analysis of ultimatum bargaining. J. Econ. Behav. Organ. 3 (4), 367–388.
Henrich, J., 2000. Does culture matter in economic behavior? ultimatum game bargaining among the machiguenga of the peruvian amazon. Am. Econ. Rev. 90 (4), 973–979. https://doi.org/10.1257/aer.90.4.973
Henrich, J., Boyd, R., 1998. The evolution of conformist transmission and the emergence of between-group differences. Evol. Hum. Behav. 19 (4), 215–241. https://doi.org/10.1016/S1090-5138(98)00018-X
Henrich, J., Boyd, R., 2001. Why people punish defectors: weak conformist transmission can stabilize costly enforcement of norms in cooperative dilemmas. J. Theor. Biol. 208 (1), 79–89. https://doi.org/10.1006/jtbi.2000.2202
Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J., 2010. Markets, religion, community size, and the evolution of fairness and punishment. Science 327 (5972), 1480–1484. https://doi.org/10.1126/science.1182238
Herrmann, B., Thöni, C., Gächter, S., 2008. Antisocial punishment across societies. Science 319 (5868), 1362–1367. https://doi.org/10.1126/science.1153808
Huck, S., Oechssler, J., 1999. The indirect evolutionary approach to explaining fair allocations. Games Econ. Behav. 28 (1), 13–24. https://doi.org/10.1006/game.1998.0691
Kandori, M., Mailath, G.J., Rob, R., 1993. Learning, mutation, and long run equilibria in games. Econometrica 61 (1), 29–56. https://doi.org/10.2307/2951777
Khan, A., 2022. Expected utility versus cumulative prospect theory in an evolutionary model of bargaining. J. Econ. Dyn. Control 137, 104332. https://doi.org/10.1016/j.jedc.2022.104332
Kirchkamp, O., Nagel, R., 2007. Naive learning and cooperation in network experiments. Games Econ. Behav. 58 (2), 269–292. https://doi.org/10.1016/j.geb.2006.04.002
Kreps, D.M., Wilson, R., 1982. Sequential equilibria. Econometrica 50 (4), 863–894.
Li, X., 2018. The Socially Embedded Individual. Ph.D. thesis. University of Nottingham.
Li, X., Molleman, L., van Dolder, D., 2021. Do descriptive social norms drive peer punishment? conditional punishment strategies and their impact on cooperation. Evol. Hum. Behav. 42 (5), 469–479. https://doi.org/10.1016/j.evolhumbehav.2021.04.002
Morris, S., Rob, R., Shin, H.S., et al., 1995. P-Dominance and belief potential. Econometrica 63 (1), 145–157. https://doi.org/10.2307/2951700
Norman, T. W.L., 2009. Rapid evolution under inertia. Games Econ. Behav. 66 (2), 865–879. https://doi.org/10.1016/j.geb.2008.10.002
Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature 359 (6398), 826–829. https://doi.org/10.1038/359826a0
Nowak, M.A., Tarnita, C.E., Antal, T., 2010. Evolutionary dynamics in structured populations. Philos. Trans. R. Soc. B: Biol. Sci. 365 (1537), 19–30. https://doi.org/10.1098/rstb.2009.0215
Oechssler, J., et al., 1997. Decentralization and the coordination problem. J. Econ. Behav. Organ. 32 (1), 119–135. https://doi.org/10.1016/S0167-2681(96)00022-4
Oechssler, J., 1999. Competition among conventions. Comput. Math. Organ. Theory 5 (1), 31–44. https://doi.org/10.1023/A:1009646410168
Ohtsuki, H., Hauert, C., Lieberman, E., Nowak, M.A., 2006. A simple rule for the evolution of cooperation on graphs and social networks. Nature 441 (7092), 502–505. https://doi.org/10.1038/nature04605
Rabin, M., 1993. Incorporating fairness into game theory and economics. Am. Econ. Rev. 83 (5), 1281–1302.
Rand, D.G., Arbesman, S., Christakis, N.A., 2011. Dynamic social networks promote cooperation in experiments with humans. Proc. Natl. Acad. Sci. 108 (48), 19193–19198. https://doi.org/10.1073/pnas.1108243108
Santos, F.C., Pacheco, J.M., 2005. Scale-free networks provide a unifying framework for the emergence of cooperation. Phys. Rev. Lett. 95 (9), 098104.
Sugden, R., 1986. The Economics of Rights, Co-Operation and Welfare. Basil Blackwell Oxford.

Sugden, R., 2000. The motivating power of expectations. Rationality, Rules, Struct., 103–129. https://doi.org/10.1007/978-94-015-9616-9_7
Szabó, G., Fáth, G., 2007. Evolutionary games on graphs. Phys. Rep. 446 (4–6), 97–216. https://doi.org/10.1016/j.physrep.2007.04.004
Tabellini, G., 2008. Presidential address: institutions and culture. J. Eur. Econ. Assoc. 6 (2–3), 255–294. https://doi.org/10.1162/JEEA.2008.6.2-3.255
Weidenholzer, S., 2010. Coordination games and local interactions: a survey of the game theoretic literature. Games 1 (4), 551–585. https://doi.org/10.3390/g1040551
Young, H.P., 1993. The evolution of conventions. Econometrica 61 (1), 57–84. https://doi.org/10.2307/2951778
Young, H.P., 1998. Individual Strategy and Social Structure: An Evolutionary Theory of Institutions. Princeton University Press.
Young, H.P., Burke, M.A., 2001. Competition and custom in economic contracts: a case study of illinois agriculture. Am. Econ. Rev. 91 (3), 559–573. https://doi.org/10.1257/aer.91.3.559