Self-Similar Beliefs in Games with Strategic Substitutes

Mengke Wang*†

November 3, 2018

Most recent version available here.

Abstract

This paper studies strategic situations where a population of heterogeneous players are randomly matched with each other to play games with strategic substitutes and players have incomplete information about their opponents' private types. If players hold type-independent beliefs about their opponents' types, then in equilibrium players' actions are monotonic with respect to their types. Since players' private types are often not observable to the analyst, to understand what kind of observable behavior can be explained by this model, a representation result is established for this model when the analyst observes how the population behaves on an aggregate level. Of course, a model with type-independent beliefs may not be justified, since types could be correlated in many applications. Moreover, in experiments where individuals are randomly matched to play games with strategic substitutes, they report systematically heterogeneous conjectures about their opponents' actions: Players who act more aggressively also conjecture that their opponents would act more aggressively. This not only contradicts the type-independent belief model, but is also counterintuitive because in games with strategic substitutes, opponents' aggressive behavior discourages players from playing aggressively. A model is then proposed where players have self-similar beliefs. It captures the intuition that higher types believe that their opponents are also of higher types and fits the experimental observations. One important and surprising result is that models with type-independent beliefs and self-similar beliefs are observationally equivalent for many payoff parameters, that is they have identical behavioral implications.

*Email address: mengke.wang@duke.edu

†I am grateful to my advisors Philipp Sadowski and Todd Sarver for their suggestions, as well as continuous support and encouragement. I also thank Atila Abdulkadiroglu, Attila Ambrus, Arjada Bardhi, Federico Bugni, Sylvain Chassang, Allan Collard-Wexler, Yuhta Ishii, Dan Kovenock, Rachel Kranton, Shengwu Li, Bart Lipman, David McAdams, Peter Norman-Sorensen, Harry Di Pei, Luca Rigotti, Curtis Taylor, Chris Timmins, Xiao Yu Wang, Daniel Xu and Huseyin Yildirim for their comments and discussions.

1 Introduction

In this paper, I study the behavioral implications of players' beliefs about each other when they engage in games with strategic substitutes. Many strategic situations of economic interest fall into this category. Take entry games, for example: Two firms simultaneously decide whether to enter a market. The market generates a higher revenue for a firm if it is the only firm that operates in it. The potential revenue from the market for one or two firms is common knowledge. However, the cost associated with entering the market is private information for each firm. The lower the cost, the more beneficial it is to enter the market. A firm's belief about its competitor's private cost therefore affects its conjecture over its competitor's action, and thus affects its expected revenue from entering the market. The higher the probability that the competitor enters the market, the more beneficial it is to stay out of the market.

I compare the behavioral implications of two modeling choices across different games, assuming that analysts can observe game parameters that are common knowledge and players' actions, but not players' private types or their conjectures about their opponents' actions. The most parsimonious model is where players hold type-independent beliefs. In other words, regardless of their own types, players hold the same belief about their opponents' type. After characterizing the equilibria, a natural nextstep is to ask first, what the model implies about players' observable behavior; second, if players' behavior is observed, whether parameters exist to explain the data; and third, if such parameters do exist, whether they are unique and how they can be identified. To address these questions, I establish a representation result for this model when analysts observe how a population behaves across games. Then I demonstrate with a constructive proof that the parameters are uniquely identified.

However, this model may not accurately capture the beliefs players hold. In the entry game example, there could be common factors that affect both players' costs to enter a market. For instance, they may be faced with similar local labor cost, rent, infrastructure cost and so on. Therefore, suppose a firm is faced with a low cost to enter a market; it makes sense for it to believe that its competitor is also faced with a relatively low entering cost. Generally speaking, suppose players hold a common prior about how types are jointly distributed and Bayesian update their beliefs about their opponents' type upon observing their own type. If types are believed to be affiliated, then as a result, higher types believe that their opponents' are more likely to be of higher types.

There is also evidence from lab experiments where individuals are asked to report their conjectures over their opponents' actions. For example, Rubinstein and Salant (2016) conduct an experiment where subjects are presented with a Hawk-Dove game[1] and instructed that they are to play this game with someone randomly chosen from among several hundred people. Among other things, the subjects are then asked which action they would choose and how likely they believe

[1] The payoff matrix of this game is presented on page 9.

their opponents would play "Hawk". They find that players who choose to play "Hawk" also hold a higher belief about their opponents playing "Hawk". This experiment demonstrates that players hold systematically different conjectures about their opponents' actions.

To incorporate these considerations, I introduce a second model where players hold self-similar beliefs: those who are of higher types also believe that their opponents are of higher types. In the Rubinstein and Salant (2016) experiment, higher types are those who are less altruistic. Since players' types are not observable to the analyst, players' degrees of altruism drive both their actions and, via self-similarity, their conjectures over their opponents' actions, explaining the positive correlation observed in the experiment.

In most empirical research, players' conjectures about their opponents' actions are not observed. Instead, the only data an analyst observes are players' chosen actions across different games either on an aggregate or individual level. One important question is whether the two models are distinguishable with such observables. I demonstrate that in games where players' equilibrium actions are monotonic with respect to their types, the behavioral implications of a self-similar belief model is qualitatively equivalent to that of a type-independent belief model. Therefore, for these games, an analyst can model players' behavior as if they hold type-independent beliefs, and use the axioms in the corresponding representation result to test the model even if beliefs are likely to be self-similar.

More formally, assume that the analyst observes a continuum population of players that are randomly matched to play symmetric 2×2 games with strategic substitutes. The payoff matrix is shown below. The two actions are labeled the high action (a_H) and the low action (a_L). The payoff parameters m and n are common knowledge for both players and are observable to the analyst. θ and θ', on the other hand, are players' private information and are not observable to the analyst.

Table 1: Game of interest

	a_L	a_H
a_L	n, n	$1-\theta, \theta'$
a_H	$\theta, 1-\theta'$	m, m

It is assumed that $m + n \leqslant 1$ and $\theta, \theta' \in [0, 1]$. This game has two important features: First, the higher a player's type, the more she prefers a_H to a_L; and second, fixing any θ (or θ', for the column player), the more likely a player conjectures her opponent to play a_H, the less attractive it is for her to play a_H. The primitive is called a *population action profile*: An analyst observes the aggregate behavior of a population across different games. In other words, for given m and n, the analyst records, for example, that 60% of the population plays the high action in that game. For a different combination of m' and n', she records that 34% of the population plays the high action.

In Section 2, I demonstrate that up to renormalization, the game above captures a wide variety of strategic situations. Examples include entry games where potential revenues are common knowledge and costs to enter are private, teamwork games where rewards are common knowledge and costs of

effort are private information, and finally lab experiments where individuals are randomly matched to play Hawk-Dove games or Prisoner's Dilemma games and their degrees of altruism are private information.

In Section 3, I study a model where players' beliefs are type-independent. More specifically, all players in the population believe that their opponents' types are randomly drawn from distribution B, and the true population type distribution is H. In many games, analysts assume that players hold correct beliefs. The model permits this but does not assume it. Instead, I leave it to the data to tell whether $B = H$ by providing a simple condition to test for correct beliefs.

When players have type-independent beliefs, the equilibrium takes the simple form of a cutoff strategy: All players whose types are higher than the cutoff type play the high action and all those whose types are lower than the cutoff type play the low action. We say that a population action profile has a *type-independent belief representation* if we can find a type distribution function H and a belief function B such that if all players in the population hold belief B, for each game (m, n), let θ^* be the equilibrium cutoff type; then the fraction of players playing the high action in the cutoff equilibrium is dictated by $1 - H(\theta^*)$: the probability that a player's type is higher than the cutoff type θ^*. For different values of m and n, the cutoff type θ^* is also different. This holds true for all games with strategic substitutes, i.e. games with $m + n \leqslant 1$. I establish that a population action profile has a type-independent belief representation if and only if it satisfies the Axioms *Dominance*, *Monotonicity*, *Continuity* and *Convexity*.

Dominance says that when m (the payoff of both playing a_H) is large enough, all players play a_H and when n (the payoff of both playing a_L) is large, all players play a_L. Monotonicity says that when m increases and n decreases, then more players play a_H. These two axioms are intuitively plausible and are supported by empirical evidence. Continuity is a technical assumption and in my model it implies that there are no atoms in the true population type distribution. Convexity can be viewed as the combination of risk neutrality and an independence condition. The requirement of risk neutrality is plausible when players are firms and when the monetary stakes are small as in lab experiments.

The methodology I adopt in this section resembles that of Daley and Sadowski (2016), who study a model of magical thinking and develop axioms for the behavior of individuals across games. I demonstrate how the axiomatic approach of analyzing behavior in strategic settings can be applied to play in the games with type-dependent payoffs described above and also how it can be generalized to heterogeneous beliefs and more general payoff structures. The benefit of the axiomatic approach is that it makes precise how behavior is different under the two models.[2]

In section 4, I allow players to hold self-similar beliefs. In other words, higher types believe that their opponents are more likely to be of higher types. With such beliefs, a cutoff equilibrium does not always exist. For example, suppose individuals are randomly matched to play a Hawk-Dove

[2]Appendix B provides a precise comparison between this model and the model of Daley and Sadowski (2016) by describing the exact differences in behavior for the two models.

game. Less altruistic players will tend to favor "Hawk" over "Dove" to a greater degree than more altruistic players. At the same time, less altruistic players who believe that others in the population are similar to themselves expect more aggressive behavior from their opponents, which discourages them from playing aggressively. Therefore, in equilibrium, it is not necessarily the case that less altruistic players play more aggressively.

Similarly, in entry games, firms with lower costs tend to favor entering the market more than higher-cost firms do. However, if they believe that their competitors also have low costs, then they conjecture that there is a higher chance that their competitors would choose to enter the market, making entry less attractive to them. Therefore, in equilibrium, it is not necessarily the case that firms with lower entry costs choose to enter the market.

I establish a necessary and sufficient condition for the existence of a cutoff equilibrium in games with strategic substitutes in this context. When players' beliefs are not too self-similar, in a way that will be formally defined in Section 4, the direct effect of private types dominates and a cutoff equilibrium exists. My result complements past research which mostly focuses on the conditions for the existence of monotone pure strategy Nash equilibria for supermodular games, where strategies are complements. For example, Athey (2001) proves that in games with incomplete information, a pure strategy Nash equilibrium exists if 1) utility functions are supermodular both in players' actions and in actions and types; and 2) players' types are affiliated. McAdams (2003) extends the result to multidimensional type spaces.

Interestingly, if players have self-similar beliefs and a cutoff equilibrium exists, then higher types play the high action, while conjecturing that their opponents' are more likely to play the high action. Since players' types are often not observable, an analyst who observes conjectures over actions will record the seemingly counterintuitive pattern that those who play the high action also believe that their opponents are more likely to play the high action. This feature of the model is consistent with the experimental evidence (Rubinstein & Salant, 2016; Charness, Rigotti, & Rustichini, 2016) that players' actions and their conjectures about their opponent's action are positively correlated even in games with strategic substitutes.

Next I show that across all games where players adopt monotone pure strategies (cutoff equilibrium), the properties of a population action profile remain qualitatively the same as if players' beliefs are type-independent. Therefore, an analyst can rationalize a population's behavior in these games as if players hold type-independent beliefs, even if there is self-similarity in their actual beliefs. Moreover, with population action data in these games, an analyst can estimate the true distribution of players' types, make inferences about the model and predict the population's behavior in hypothetical games[3] as if players hold type-independent beliefs. Of course, this is limited to the case where only players' actions are observed. If in addition to actions, the analyst also ob-

[3]To make sure that predictions are made within the set of cutoff equilibria games, individual action data can be collected to identify players' unobservable private type and thus the set of games that are played with monotone pure strategies.

serves players' conjectures about their opponent's action, then the two models have very different implications on the observables. Type-independent beliefs imply that players' equilibrium conjectures are homogeneous, while self-similar beliefs imply that in cutoff equilibria players' equilibrium conjectures are positively aligned with their actions.

Finally, I also propose a definition for one population being more self-similar than another and discuss its behavioral content. Intuitively, a population is more self-similar if players' beliefs depend more on their types and their self-perceived quantiles are more dispersed. Now that players' beliefs are more dependent on their types, cutoff equilibria cease to exist in more games. In terms of behavioral implications, for a more self-similar population, the set of games where the population's behavior can be rationalized by type-independent beliefs shrinks and within this set of games, players act more similarly: they tend to side more on the dominant action, whether it is the high action or the low action depending on the strategic situation they are engaged in.

It is left unanswered in this paper whether self-similar beliefs can be sustained when players engage in similar strategic environments repeatedly. Frick, Iijima, and Ishii (2018) discuss this topic by assuming that players interact with each other assortatively: higher types are more likely to interact with higher types. They show that a unique form of misperception persists in all environments: all players believe that the people they interact with constitute a representative sample of the total population.

2 Examples

I now demonstrate that up to renormalization, the game proposed earlier captures the strategic content of a wide variety of strategic situations. Recall that the normalized 2×2 game of strategic substitutes takes the following form.

Table 2: Game of interest

	a_L	a_H
a_L	n, n	$1-\theta, \theta'$
a_H	$\theta, 1-\theta'$	m, m

Let (m, n) denote the above game. In this game, suppose a type-θ player conjectures that with probability μ her opponent plays a_H; then she weakly prefers a_H if and only if

$$\mu \cdot m + (1-\mu) \cdot \theta \geqslant \mu \cdot (1-\theta) + (1-\mu) \cdot n$$

Example 1 (Entry games)**.** The payoff matrix of a symmetric entry game is shown as below.

Table 3: Entry game

	Out	Enter
Out	$0, 0$	$0, R_H - c'$
Enter	$R_H - c, 0$	$R_L - c, R_L - c'$

R_H denotes the revenue a firm can earn when it is the only producer in the market, while R_L denotes its revenue if it is one of two producers in the market. $R_H \geqslant R_L$. $c, c' \in [0, 1]$ denote the two firms' private costs of entering the market. Suppose a firm with private cost c conjectures that its opponent enters the market with probability μ; then it weakly prefers to enter the market if and only if

$$\mu \cdot (R_L - c) + (1 - \mu) \cdot (R_H - c) \geqslant 0.$$

In this context, let a_H denote the action of entering the market and a_L denote the action of staying out of the market. To show that this game is strategically equivalent to the game in Table 2, let $\theta = 1 - c$, $\theta' = 1 - c'$, $m = R_L$ and $n = 1 - R_H$. Then

$$\begin{aligned} & \mu \cdot (R_L - c) + (1 - \mu) \cdot (R_H - c) \geqslant 0 \\ \iff & \mu \cdot R_L + (1 - \mu) \cdot (1 - c) \geqslant \mu \cdot c + (1 - \mu)(1 - R_H) \\ \iff & \mu \cdot m + (1 - \mu) \cdot \theta \geqslant \mu \cdot (1 - \theta) + (1 - \mu) \cdot n. \end{aligned}$$

Moreover, $m + n \leqslant 1 \iff R_L \leqslant R_H$.

Intuitively, type θ describes how cost efficient a firm is. m and n measure the desirability of the outcomes where both firms choose to enter and where both firms choose to stay out of the market respectively. When R_H is higher, the firms forgo a larger amount of revenue by staying out of the market. When R_L is higher, entering the market is more attractive even if the other firm is also entering the market.

Example 2 (Teamwork games)**.** Two teammates collaborate on a project. They each can choose to exert effort or shirk on the project.

Let w_1 and w_2 denote the commonly known payoffs both players get if one or two players work on the project, and $c, c' \in [0, 1]$ denote the two players' private costs of exerting effort. The payoff if neither player works on the project is normalized to 0.

Table 4: Team effort game

	Work	Shirk
Work	$w_2 - c, w_2 - c'$	$w_1 - c, w_1$
Shirk	$w_1, w_1 - c'$	$0, 0$

If a player whose cost of effort is c conjectures that with probability μ her teammate shirks, and with probability $1-\mu$ her teammate works on the project, then she weakly prefers to shirk if and only if

$$\begin{aligned} &\mu \cdot 0 + (1-\mu) \cdot w_1 \geqslant \mu \cdot (w_1 - c) + (1-\mu) \cdot (w_2 - c) \\ \iff &\mu \cdot m + (1-\mu) \cdot \theta \geqslant \mu \cdot (1-\theta) + (1-\mu) \cdot n, \end{aligned}$$

where $\theta = c$, $n = w_2 - w_1$ and $m = 1 - w_1$.

Intuitively, let "Shirk" be the high action and "Work" be the low action. θ measures a player's cost to exert effort. Therefore, the higher a player's type is, the more benefit she gets from shirking. m measures the desirability of the outcome where both players shirk. The higher w_1 is, the less desirable this outcome is.

Moreover, with this renormalization,

$$m + n \leqslant 1 \iff w_2 - w_1 \leqslant w_1.$$

In words, the marginal benefit of more people working on the project decreases. This is the key property that generates the strategic substitution effect in these games.

Example 3 (Lab experiments with monetary payoffs and altruistic individuals)**.** A common setup in lab experiments is as follows: individuals are presented with the payoff matrix of a game and instructed that they are going to be randomly matched with some other participant to play this game. Analysts observe players' actions and sometimes also their reported conjectures on how their opponents' would play. In this case, monetary outcomes for both players are explicitly specified and thus can be viewed as common knowledge. However, as experimental research shows, individual players can also care about their opponents' payoffs. To what degrees they do so is their private information.

Suppose players have altruistic preferences: Let π_i and π_j be the payoffs for player i and her opponent player j; then player i's utility is $U_i(\pi_i, \pi_j, \theta_i) = \theta_i \pi_i + (1-\theta_i)\pi_j$, where $\theta_i \in \Theta := [0,1]$ is player i's private type. It is easy to see that the game of interest depicts the strategic situation where the row player is of type θ, the column player is of type θ' and they are matched to play the following game:

Table 5: Symmetric 2×2 game with monetary payoffs

	a_L	a_H
a_L	n, n	$0, 1$
a_H	$1, 0$	m, m

More generally, suppose players are randomly matched to play any symmetric 2×2 game, with the payoff matrix shown below, where $x > y$ and $k + l \leqslant x + y$.

Table 6: General symmetric 2×2 game with monetary payoffs

	a_L	a_H
a_L	k, k	y, x
a_H	x, y	l, l

By assuming $x > y$, we label the better-off player's action as a_H when two players' actions are different. It is almost without loss of generality except that the case where both players' payoffs are the same is ruled out. On the other hand, the condition $k+l \leqslant x+y$ implies that $x-k \leqslant l-y$, the left hand side of which is the monetary gain of a player switching from a_L to a_H when her opponent plays a_L, while the right hand side is that when her opponent plays a_H. In other words, these are games with strategic substitutes.

The games used by Rubinstein and Salant (2016) fall into this category, where a_H (or action B in their experiments) is interpreted as "Hawk" in a Hawk-Dove game and a_L (or action A in their experiments) is "Dove".

Table 7: Rubinstein Salant experiment game

	A	B
A	$30, 30$	$30, 70$
B	$70, 30$	$0, 0$

The Prisoner's Dilemma (PD) games used by Charness et al. (2016) also satisfy this condition. In these games, the high action a_H corresponds to "Defect" and the low action a_L corresponds to "Cooperate". Charness et al. (2016) use different values for k, namely $k = 3, 4, 5, 6$. For all these values, the monetary gain of switching from "Cooperate" to "Defect" is greater when the opponent cooperates.

Table 8: Charness et al. experiment game

	Cooperate	Defect
Cooperate	k, k	$1, 7$
Defect	$7, 1$	$2, 2$

Therefore, using the labels in Table 6 a type-θ player prefers a_H if and only if

$$\mu \cdot l + (1 - \mu) \cdot [y + (x - y)\theta] \geqslant \mu \cdot [x + (y - x)\theta] + (1 - \mu) \cdot k,$$

where μ is the probability with which a player conjectures that her opponent plays a_H. Let

$$m = \frac{l-y}{x-y} \text{ and } n = \frac{k-y}{x-y},$$

then the condition is equivalent to

$$\mu \cdot m + (1-\mu) \cdot \theta \geqslant \mu \cdot (1-\theta) + (1-\mu) \cdot n.$$

As a final note, the topics of other-regarding preferences and self-reported beliefs have received a lot of attention from both theorists and experimental economists. For example, Charness and Rabin (2002) use a series of games to figure out the appropriate form of social preferences. Costa-Gomes and Weizsäcker (2008) discuss whether the elicited beliefs are the true beliefs players best respond to. They assume that all players hold the same belief and propose a statistical model to test and later reject the null hypothesis that the underlying beliefs for actions and the reported beliefs coincide. What is lacking in this literature is a self-contained model to properly deal with the heterogeneity of players in terms of both payoff relevant types and interim beliefs to match the choice data collected by experimental economists. The common practice in the existing literature is to credit the heterogeneity entirely to noise. My work contributes to filling this gap by laying out an economic model, upon which more sophisticated statistical models can be built.

3 Type-independent beliefs and population action profiles

A continuous population I of players are randomly matched to play the following symmetric 2×2 game. Let $X := \{a_L, a_H\}$ denote the set of actions. The values of m and n are common knowledge for both players and also observable to the analyst. Let $\mathcal{G} := \{(m,n) \mid m+n \leqslant 1\}$ denote the set of all games with strategic substitutes. $\theta \in \Theta := [0,1]$ is the row player's private type that is unobservable to both her opponent and the analyst. Similarly, $\theta' \in [0,1]$ is the column player's private type that is unobservable to the row player and the analyst.

Table 9: Game of interest

	a_L	a_H
a_L	n, n	$1-\theta, \theta'$
a_H	$\theta, 1-\theta'$	m, m

In this section, I study the parsimonious model where players hold type-independent beliefs. In other words, regardless of their private type, all players hold belief B about their opponents' types. B is a continuous cumulative density function (c.d.f.) on Θ. For any game $g \in \mathcal{G}$, let $u_g(a, a', \theta)$ denote a type-θ player's payoff if she plays action a and her opponent plays a'.

Definition 1. For any game $g \in \mathcal{G}$ and belief B, a *subjective Bayesian Nash equilibrium* (SBNE) is a mapping $\sigma : \Theta \to \Delta(X)$ such that, for $\mu(a) := \int_{\theta' \in \Theta} \sigma'_\theta(a) \mathrm{d}B(\theta'), \forall a \in X$, and for $\theta \in \Theta$ and any $a, a' \in X$ with $\sigma_\theta(a) > 0$,

$$\mu(a_H)u_g(a, a_H, \theta) + \mu(a_L)u_g(a, a_L, \theta) \geqslant \mu(a_H)u_g(a', a_H, \theta) + \mu(a_L)u_g(a', a_L, \theta).$$

There are two possible interpretations of this equilibrium concept: First, the equilibrium strategy is the symmetric equilibrium strategy adopted by both players in the game. In other words, the equilibrium strategy is a contingent plan of actions for both players before their types (preferences and beliefs) are determined by nature.

The second interpretation, which better fits the setup in this paper, is that equilibrium is a description of the population behavior; each player has a private preference type and also holds a belief about others' types. The equilibrium describes a situation where given every type's choice of action, players form a common conjecture over actions based on their common belief about others' types, and in equilibrium all types best respond to this common conjecture over actions.

Definition 2. An SBNE (σ, μ) is a *cutoff equilibrium* if there exists $\theta^* \in \Theta$ such that

$$\sigma_\theta = a_L, \ \forall \theta < \theta^*; \quad \sigma_\theta = a_H, \ \forall \theta > \theta^*.$$

When players' beliefs are type-independent, a unique equilibrium always exists and takes the form of a cutoff equilibrium.

Proposition 1. *For any belief B and any game $g \in \mathcal{G}$, a unique SBNE (σ, μ) exists. Moreover, the unique SBNE is a cutoff equilibrium.*

Suppose an analyst observes how a population behaves on an aggregate level in all games $g \in \mathcal{G}$. Specifically, for each game she records what fraction of the population plays the high action. A *population action profile* is a function $A : \mathcal{G} \to [0, 1]$. For example, $A(m, n) = 0.8$ means that in game (m, n), 80% of the population play the high action.

For instance, for entry games, an analyst observes how the entry rate changes when the potential revenues (R_H and R_L) change, assuming firms' costs have a stable distribution. For the teamwork games, roughly speaking, an analyst observes how a population's overall effort level (e.g. the percentage of students who exert effort on a team project) changes when w_1 and w_2 (e.g. the grading system for the team project) changes. In lab experiments, the analyst observes what fractions of the population play a certain action given different monetary payoffs.

The next definition connects $A(m, n)$ to a model with type-independent beliefs.

Definition 3. We say that a population action profile A has a *type-independent belief representation* if there exist functions $H, B : [0, 1] \to [0, 1]$ such that

1. Function H is continuous and strictly increasing with $H(0) = 0$ and $H(1) = 1$.
2. Function B is continuous and weakly increasing with $B(0) = 0$ and $B(1) = 1$.
3. For any game (m, n), $A(m, n) = 1 - H(\theta^*)$, where θ^* satisfies

$$\theta > \theta^* \implies \theta B(\theta^*) + m(1 - B(\theta^*)) \geqslant nB(\theta^*) + (1 - \theta)(1 - B(\theta^*));$$
$$\theta < \theta^* \implies \theta B(\theta^*) + m(1 - B(\theta^*)) \leqslant nB(\theta^*) + (1 - \theta)(1 - B(\theta^*)).$$

In words, a population action profile A has a type-independent belief representation if there exist a population type distribution H and a belief distribution B such that if all players in the population hold belief B, then the fraction of players playing the high action is dictated by $1-H(\theta^*)$: the probability that a player's type is higher than the cutoff type.

The following theorem states the necessary and sufficient condition for a population action profile to have a type-independent belief representation.

Theorem 1. *A population action profile A has a type-independent belief representation if and only if it satisfies*

1. *(Dominance) $A(m, n) = 1$ for all $m \geqslant 1$ and $A(m, n) = 0$ for all $n \geqslant 1$.*
2. *(Monotonicity) Take any two games (m, n) and (m', n') such that $m, m', n, n' \leqslant 1$. If $m < m'$ and $n > n'$ then $A(m, n) < A(m', n')$.*
3. *(Continuity) For any fixed $c \leqslant 1$, function $A(c - n, n)$ is continuous in n.*
4. *(Convexity) Take any two games (m, n) and (m', n') such that $m, m', n, n' \leqslant 1$. If $A(m, n) = A(m', n') = \alpha$, then*

$$A(\gamma m + (1 - \gamma)m', \gamma n + (1 - \gamma)n') = \alpha, \ \forall \gamma \in [0, 1].$$

This theorem establishes the exact behavioral implications of a model where players hold type-independent beliefs. Therefore, one can test the model or assess its empirical plausibility based on the four axioms.

To interpret Dominance, note that when $m + n \leqslant 1$, $m \geqslant 1$ implies that $n \leqslant 0$. a_H is the dominant strategy for all types. Similarly, $n \geqslant 1$ implies that $m \leqslant 0$; and a_L is the dominant strategy for all types.

Table 10: Game of interest

	a_L	a_H
a_L	n, n	$1 - \theta, \theta'$
a_H	$\theta, 1 - \theta'$	m, m

In the lab experiment interpretation, when $m \geqslant 1$, independent of the opponent's action, playing a_H weakly improves both players' payoffs. Dominance requires that for such games, the whole population plays a_H. This axiom is very plausible in contexts where players care about each others' payoffs, but rules out strong cases of inequality aversion or negative emotions such as spite. For example, it rules out that a positive fraction of the population values the equality of the payoff distribution to the extent that they would prefer to lower both payoffs in order to achieve equality. In other words, this axiom puts restrictions on players' social preferences, if they have any.

In the context of entry games, after renormalization, Dominance says that when the revenue of both firms operating is high enough, all firms choose to enter; on the other hand, when the revenue of being the sole producer in the market is low enough, all firms choose to stay out of the market. In this context, Dominance implies that firms' costs of entry are bounded. Similarly, in the context of teamwork games, Dominance says that if the reward for being the only person working on the project is low enough, then no one works on the project. On the other hand, if the additional reward for a second person to work on the project is high enough, then everyone works on the project.

Monotonicity captures a basic sensitivity to the payoffs in the game: When m increases and n decreases, payoffs are higher for both players if they both play a_H and lower if they both play a_L. Monotonicity requires that more players play a_H in this case. For example, Charness et al. (2016) document that in their prisoner's dilemma games, when k increases, more subjects cooperate.

Table 11: Charness et al. experiment game

	Cooperate	Defect
Cooperate	k, k	$1, 7$
Defect	$7, 1$	$2, 2$

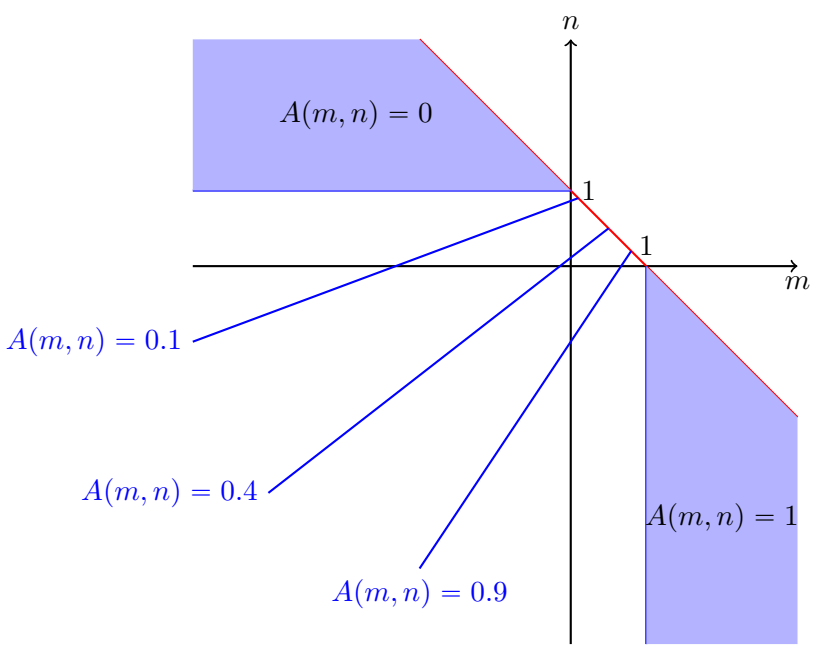
As usual, Continuity is a technical assumption that is hard to falsify. It requires that when two games are close to each other, the population's action profiles are also close. Of course, in a game theoretic context when equilibria change, continuity is not plausible without heterogeneous types. In my model, Continuity implies that there are no atoms in the true distribution of types in the population.

Convexity says that if the same fraction of the population plays the high action in two games, then that same fraction also plays the high action in all games where the payoffs are a linear combination of the payoffs in the two original games. Consider first the hypothetical strategic situation where players must decide which action to take, and afterwards with probability γ, game (m, n) is played, and with probability $1 - \gamma$, game (m', n') is played. Suppose for each player the preferred actions are the same for the two games. In other words, those who prefer a_H in one game also prefer a_H in the other game. Similarly for action a_L. Then in the new strategic situation, it makes sense that all players stick to the action they preferred in the original games. The new

strategic situation is pay-off equivalent to games where the payoffs are a linear combination of the original games if players are risk neutral and therefore their preferences over compound lotteries which involve uncertainty about nature (which game is played) and opponents' behavior can be reduced to a comparison of expected payoffs.

The following graph illustrates a typical population action profile that satisfies all four axioms. Each game (m, n) is a point on the m-n plane. The iso-value curves represent games where the same fraction of the population play a_H. For example, all the points on the iso-value curve $A(m, n) = 0.1$ represent games where 10% of the population plays a_H. The Dominance axiom corresponds to the two dominance regions where $A(m, n) = 0$ or 1. The Monotonicity and Convexity axioms imply that the iso-value curves are downward sloping straight lines that fan out in the non-dominance region.

Figure 1: Iso-value curves for A on m-n plain



Axioms 1, 2 and 4 constitute testable implications for the type-independent belief model. For example, suppose an analyst observes that in both games g_1 and g_2, 20% of the population play a_H. She can test the Convexity axiom by looking at games with payoffs that are linear combinations of those of g_1 and g_2. If 20% of the population play a_H also in these games, then the Convexity axiom is not rejected. Otherwise, the axiom and therefore the model itself is rejected.

Sketch of proof of Theorem 1. It is easy to check that Dominance is satisfied if a population action profile has a type-independent belief representation. In the non-dominance region, for any game

(m, n) the cutoff type θ^* is indifferent between a_L and a_H if and only if

$$(1 - m - n)B(\theta^*) = 1 - m - \theta^*.$$

Analysis of the solutions to this equation implies the rest of the axioms.

In the following steps, I demonstrate how functions H and B are constructed. In games $(1-n, n)$ where $n \in [0, 1]$, the cutoff type $\theta^* = n$. Therefore, it has to be the case that $H(n) := 1-A(1-n, n)$ for all $n \in [0, 1]$. Continuity and Monotonicity imply that function H is continuous and strictly increasing. Take any other game (m', n') where $m'+n' < 1$ and $A(m', n') = A(1-n, n)$. Continuity implies that such a game always exist. The key intuition is that since the same fraction of the population plays a_H, it has to be the case that in game (m', n') the cutoff type is also n. Therefore, let

$$B(n) := \frac{1 - m' - n}{1 - m' - n'} \implies (1 - m' - n')B(n) = 1 - m' - n.$$

The axioms Continuity, Monotonicity and Convexity imply that function B is well-defined, continuous and weakly increasing.

From the construction of H and B, it is intuitive that given any population action profile that has a type-independent belief representation, both parameters are uniquely identified. Moreover, a limited set of games need to be investigated in order to identify these parameters. For example, after identifying function H with games where $m + n = 1$, an analyst can then use games with $m + n = c$ to identify function B, where c is a real number strictly less than 1. Moreover, she can restrict attention to games with $m, n \leqslant 1$ since otherwise there is a strictly dominant strategy for all types.

Proposition 2. *Suppose a population action profile A has shared belief representations (H, B) and (H', B'), then $H = H'$ and $B = B'$.*

Comparative statics

If a population has more higher types or all players in the population believe that their opponents are of lower types than those in another population do, then higher fractions of the former population play the high action in all games.

Proposition 3. *Suppose population action profiles A and A' have shared belief representations (H, B) and (H', B'). If $H(\theta) \geqslant H'(\theta)$ and $B(\theta) \leqslant B'(\theta)$ for all $\theta \in [0, 1]$, then $A(m, n) \leqslant A'(m, n)$ for all (m, n).*

One can further disentangle the effects of preference shift and belief shift. The intuition is that whether two games have the same cutoff type depends solely on players' beliefs B about the population type distribution. On the other hand, fix the cutoff type, what fraction of the population

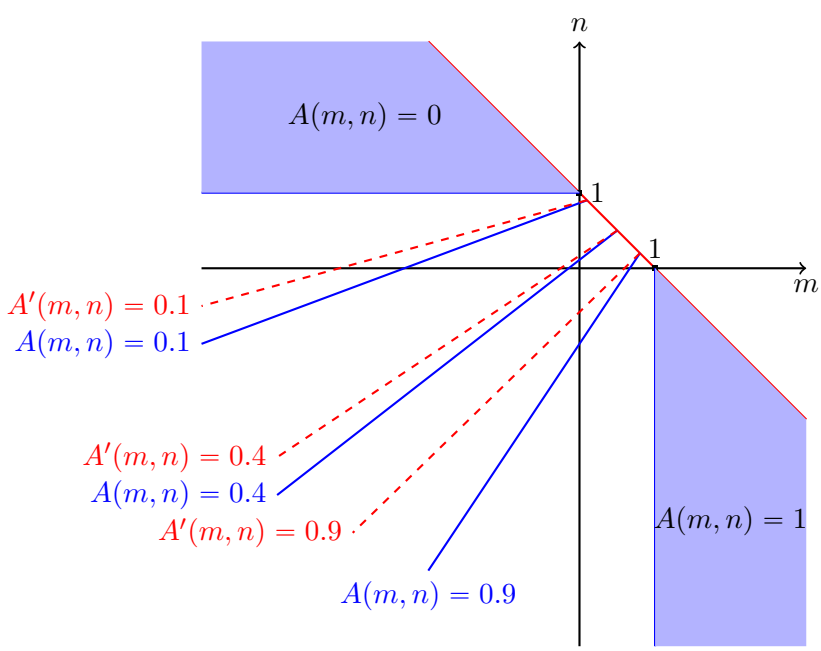
plays the high action depends only on the true population type distribution H. Since the cutoff type for a game with $m + n = 1$ does not depend on players' beliefs, the fraction of the population playing a_H in this game is determined by H. However, which games are played the same way, i.e. the same fraction of the population play a_H, is determined by B.

Suppose two populations have the same population type distribution but one population believes that it consists of lower type players. In addition to a higher fraction of the high action being played in the low belief population, she also expects to observe that for games where $m + n = 1$, the two populations behave the same since beliefs does not play a role in these games. On the other hand, if she observes two populations behave this way, then it has to be the case that they have the same type distribution, and the behaviorally more aggressive population holds a lower belief.

Proposition 4. *Suppose population action profiles A and A' have shared belief representations (H, B) and (H', B'). $H(\theta) = H'(\theta)$ and $B(\theta) \leqslant B'(\theta)$ for all $\theta \in [0, 1]$ if and only if $A(m, n) \leqslant A'(m, n)$ for all (m, n) and $A(m, n) = A'(m, n)$ when $m + n = 1$.*

On the m-n plane, all the iso-value curves rotate clock-wise around their intersections with the boundary.

Figure 2: Iso-value curves when belief shifts downwards



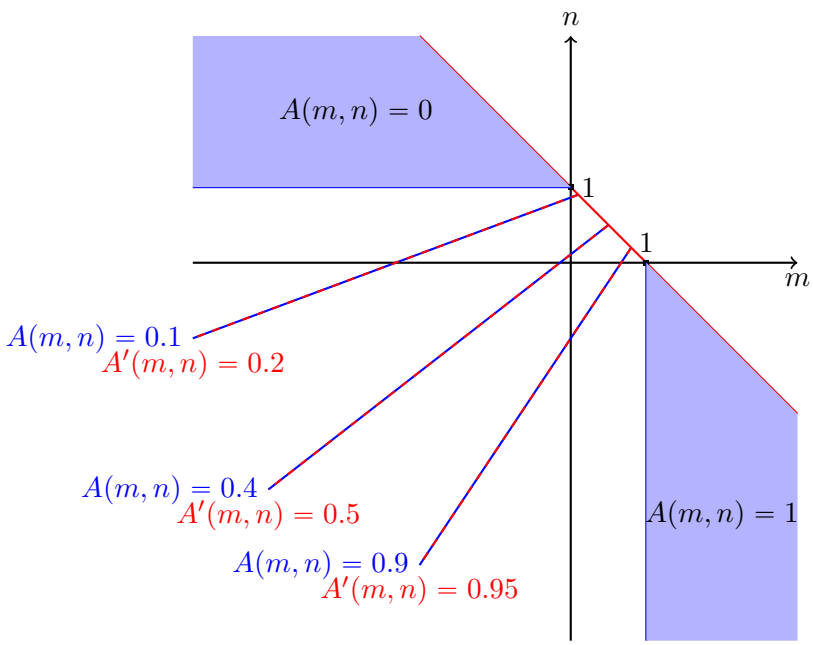
Suppose two populations share the same belief, but differ in terms of population type distributions. If two games are played the same by one population, then they are also played the same by the other population; except for that more people play a_H for both games by the population that has more higher types. The reverse is also true.

Proposition 5. *Suppose population action profiles A and A' have shared belief representations (H, B) and (H', B'). $H(\theta) \geqslant H'(\theta)$ and $B(\theta) = B'(\theta)$ for all $\theta \in [0, 1]$ if and only if*

1. *$A'(m, n) \geqslant A(m, n)$ for all $(m, n) \in \mathcal{G}$, and*
2. *$A(m, n) = A(m', n') \implies A'(m, n) = A'(m', n')$, for all $(m, n), (m', n') \in \mathcal{G}$.*

On the m-n plain, the two sets of iso-value curves overlap with each other, except that all the curves are of higher values, as is shown in the graph below.

Figure 3: Iso-value curves when types shift upwards



Now that both the population type distribution and shared belief can be identified, a natural question to ask is how these two parameters compare. The following theorem states the necessary and sufficient condition for them to coincide.

Theorem 2. *Suppose a population action profile A has shared belief representation (H, B). $H(\theta) = B(\theta)$ for all θ if and only if*

$$\begin{aligned} & A(m, n) = \alpha \in (0, 1) \\ \implies & A(m + (1 - \alpha)\delta, n + \alpha\delta) = \alpha, \forall \delta \leqslant 1 - m - n. \end{aligned}$$

In words, to test whether the players in a population share the *correct* belief about the population type distribution, an analyst can pick any game (m, n) and change the values of m and n by amounts in the ratio of $(1 - \alpha) : \alpha$, where α is the fraction of the population playing a_H in

game (m, n). Suppose players hold the correct belief; then in the new game, the analyst should also observe that α of the population play a_H.

The intuition is as follows: Suppose a type-θ row player conjectures that her opponent plays a_H with probability μ. If the values of m and n are changed by the ratio of $(1-\mu) : \mu$, then the player's expected utilities for playing the two actions are changed by the same amount, which implies that her preference over the two actions remains unchanged.

Table 12: Type-θ row player's utilities

	a_L: $1-\mu$	a_H: μ
a_L	n	$1-\theta$
a_H	θ	m

$\rightarrow$

	a_L: $1-\mu$	a_H: μ
a_L	$n + \mu\delta$	$1-\theta$
a_H	θ	$m + (1-\mu)\delta$

Suppose players hold the correct belief about their opponents' types; then their conjectures are true. Take any game where α of the population play a_H; players should also conjecture that with probability α their opponent would play a_H. If payoffs are changed in the ratio of $(1-\alpha) : \alpha$, players' preferences over the two actions do not change. Therefore, α of the population play a_H in the new game.

4 Self-similar beliefs and individual action profiles

In this section, I study a model where players have self-similar beliefs. In terms of observables, if players' conjectures about their opponents' actions are directly elicited, as in the lab experiments previously mentioned, the two models make very distinct predictions: A type-independent belief model implies that players' reported conjectures are homogeneous, while a self-similar belief model implies otherwise. For most empirical applications, where players' conjectures are not observed, I demonstrate that a self-similar belief model predicts the same qualitative behavioral patterns as a type-independent model does for many games. In other words, for many games, the more parsimonious model of type-independent beliefs provides an accurate description of players' behavior even if players do in fact hold self-similar beliefs.

4.1 Self-similar beliefs

Recall the strategic situation: Players are randomly matched with each other to engage in a symmetric 2×2 game with strategic substitutes, $g \in \mathcal{G}$. Players have private information about their own payoff relevant type. Moreover, players hold type-dependent beliefs about each others' types. Fix any $\theta \in [0, 1]$, let $F(\cdot; \theta)$ denote the cumulative density function (c.d.f.) of the belief of a type-θ player. Then the *belief profile* of population I is fully described by $F : \Theta \times \Theta \rightarrow [0, 1]$, where $\Theta = [0, 1]$ and $F(\theta'; \theta)$ denotes the probability a type-θ player assigns to interval $[0, \theta']$.

Definition 4. A belief profile F is *continuous* if the function $F : \Theta \times \Theta \to [0, 1]$ is continuous on $\Theta \times \Theta$.

The definition captures continuity in two ways. First, no probability mass is assigned to any type by any individual. Second, when two individuals' types are close to each other, their beliefs about the population type distribution are also close.

Definition 5. A belief profile F is *self-similar* if for any $\theta \leqslant \theta'$ and $\theta^* \in \Theta$, $F(\theta^*; \theta) \geqslant F(\theta^*; \theta')$.

Definition 6. A belief profile F is *consistent* if $F(\theta; \theta)$ is increasing in θ.

In words, a population has a self-similar belief profile if higher types' beliefs first order stochastic dominate (FOSD) lower types' beliefs. Mathematically, $F(\cdot; \cdot)$ is increasing in the first argument and decreasing in the second. In addition, players' beliefs are consistent if they do have a rough sense as to where they stand in the population: higher types believe they are ranked higher in the population than lower types believe they are.

As is previously mentioned, the self-similar interim beliefs could be a result of a standard Bayesian update from a common prior when players' types are affiliated. Players observe their own types and use this information to form posterior beliefs about the type distribution in the population. Consider the following simple example with discreet types.

Example 4 (Self-similar beliefs generated by a Bayesian update)**.** Suppose there are two possible states of the world: In the high state s_H, 2/3 of the population are of high type θ_H and 1/3 of the population are of low type θ_L. In the low state s_L, 1/3 of the population are of high type θ_H and 2/3 of the population are of low type θ_L. A priori, all players assign equal probability to s_H and s_L.

Once players' private types are revealed, they Bayesian update their beliefs about the state of the world based on their own types. For players of high type θ_H:

$$\text{Pr.}(s = s_H) = \frac{1/2 \cdot 2/3}{1/2 \cdot 2/3 + 1/2 \cdot 1/3} = \frac{2}{3}; \quad \text{Pr.}(s = s_L) = \frac{1/2 \cdot 1/3}{1/2 \cdot 2/3 + 1/2 \cdot 1/3} = \frac{1}{3}.$$

Therefore, a high type player believes that

$$\text{Pr.}(\theta = \theta_H) = \frac{2}{3} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{1}{3} = \frac{5}{9}; \quad \text{Pr.}(\theta = \theta_H) = \frac{2}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} = \frac{4}{9}.$$

Similarly, players of low type θ_L believe that with probability 4/9 her opponent is of high type and with probability 5/9 her opponent is of low type. Thus self-similar beliefs, where higher types believe that there are more high types and lower types believe that there are more low types, can be the result of a Bayesian update.

The false-consensus literature (see Marks and Miller (1987) for a survey) in psychology also documents and provides explanations for the phenomenon that individuals tend to perceive the

general population as more similar to themselves than it actually is. In this paper, I abstract away from how players form their beliefs about each other. Instead, I focus on the implications of self-similar beliefs for players' observable behavior.

4.2 Cutoff equilibrium

The subjective Bayesian Nash equilibrium (SBNE) is generalized as follows to incorporate type-dependent beliefs. In words, players form conjectures about their opponents' actions based on 1) their type-dependent beliefs about their opponents' types and 2) how different types play in equilibrium. An SBNE specifies how each type plays so that all types best respond to their own beliefs. For any game $g \in \mathcal{G}$, let $u_g(a, a', \theta)$ denote a type-θ player's payoff if she plays action a and her opponent plays a'.

Definition 7. For any game $g \in \mathcal{G}$ and belief profile F, a subjective Bayesian Nash equilibrium is a mapping $\sigma : \Theta \to \Delta(X)$ such that, for $\mu_\theta(a) := \int_{\theta' \in \Theta} \sigma_{\theta'}(a) \mathrm{d}F(\theta'; \theta)$, and for any $a' \in X$, $\theta \in \Theta$ and $\sigma_\theta(a) \neq 0$,

$$\mu_\theta(a_H)u_g(a, a_H, \theta) + \mu_\theta(a_L)u_g(a, a_L, \theta) \geqslant \mu_\theta(a_H)u_g(a', a_H, \theta) + \mu_\theta(a_L)u_g(a', a_L, \theta).$$

There may be multiple equilibria for $g \in \mathcal{G}$ and F. I propose to the previously cited evidence that beliefs and actions are positively aligned as a criterion to select among those equilibria: The only pure strategy equilibria that generate the positive alignment between actions and beliefs are cutoff equilibria. This statement is formalized in the following claim.

Claim 1. *Let σ be a pure strategy SBNE for game $g \in \mathcal{G}$ and belief profile F. Let μ be the equilibrium conjecture. Suppose for any $\theta, \theta' \in \Theta$, $\sigma_\theta > \sigma_{\theta'} \implies \mu_\theta > \mu_{\theta'}$. Then for any $t \leqslant t' \in \Theta$,*

$$\sigma_t \leqslant \sigma_{t'}$$

The intuition is as follows. Pick any two players; suppose the player who plays the higher action also holds a higher conjecture about her opponent playing the high action. Since strategies are substitutes, the fact that both players best respond to their conjectures implies that the player who plays the high action must be of a higher type.

However, a cutoff equilibrium does not always exist. In particular, when players' beliefs are self-similar, players' preferences and beliefs affect the rankings of the two actions in different directions. On the preference side, higher types prefer the high action to a greater extent. On the belief side, however, they also conjecture that, since there are more high types like themselves, more players in the population will play the high action. In games with strategic substitutes, such conjectures discourage them from playing the high action. Depending on which of the two effects dominates, it may or may not be the case that higher types play the high action more often in equilibrium.

For instance, as a later example shows, it could be the case that in equilibrium, the highest types and the lowest types play a_L, while those in the middle play a_H. (See the example on page 30.)

The following theorem establishes the necessary and sufficient condition for a unique cutoff equilibrium to exist for the class of games with $m + n = c$.

Theorem 3. *For game $g \in \mathcal{G}$ and belief profile F that is continuous and consistent, a unique cutoff equilibrium exists for all games (m, n) where $m + n = c < 1$ if and only if for any $\theta, \theta' \in [0, 1]$,*

$$-\frac{F(\theta';\theta) - F(\theta';\theta')}{\theta - \theta'} \leqslant \frac{1}{1-c}.$$

Note that the right hand side (RHS) is always positive. If the left hand side (LHS) is negative, then the inequality holds. This is the case where higher types believe there are fewer higher types:

$$-\frac{F(\theta';\theta) - F(\theta';\theta')}{\theta - \theta'} \leqslant 0.$$

In this case, the preference effect and the belief effect work in the same direction. A cutoff equilibrium always exists.

When the LHS is nonnegative, the preference effect and the belief effect work in opposite directions. The condition can be rewritten as follows:

$$\underbrace{\left|F(\theta';\theta) - F(\theta';\theta')\right|}_{\text{Change in Beliefs}} \cdot \underbrace{(1-c)}_{\text{Substitution effect}} \leqslant \underbrace{\left|\theta - \theta'\right|}_{\text{Preference Effect}}.$$

It simply reads that the preference effect dominates the belief effect. The right hand side of the inequality describes how much more inclined towards the high action a higher type is compared with a lower type. The left hand side is the product of the change in beliefs, a population-specific measure, and the strength of the substitution effect, a game-specific measure.

The preference effect and change in beliefs are both self-explanatory. The measure of the substitution effect justifies more clarification. For any game (m, n), $1 - c = 1 - m - n$. Therefore, the lower the values of m and n, the stronger the substitution effect. In the context of entry games, for example, $m = R_L$ and $n = 1 - R_H$. Low values of m and n imply that the revenue of being the sole producer in the market is especially high, and the revenue of being one of two producers in the market is especially low. In this situation, firms have strong incentives to stay out of the market when it is conjectured that their opponents would enter the market. Conversely, firms have strong incentives to enter the market when it is conjectured that their opponents would stay out of the market. This is then a game of strong strategic substitutes. A similar intuition applies to other applications.

As m and n decrease, the substitution effect becomes stronger, and hence players' preferences over the two actions become more sensitive to their conjectures over their opponents' actions. This

amplifies the belief effect. When the belief effect dominates, a cutoff equilibrium ceases to exist. The corollary below states that if a cutoff equilibrium exists for games with stronger strategic substitutes then it must also exist for games with weaker strategic substitutes.

Corollary 1. *For any belief profile F that is continuous, consistent and self-similar and any $c < 1$, if a unique cutoff equilibrium exists for all games (m, n) where $m + n = c$, then a unique cutoff equilibrium exists for all games (m', n') where $m' + n' \in (c, 1)$.*

Similarly, suppose in one population, players' beliefs depend more on their type. Then, for any fixed game, the belief effect is also stronger for this population. When the belief effect dominates, a cutoff equilibrium ceases to exist. The corollary below states that if a cutoff equilibrium exists when beliefs depend more on the own type, then one must also exist for less dependence.

Corollary 2. *Suppose belief profiles F and F' are continuous, consistent, self-similar and*

$$|F(\theta'; \theta) - F(\theta'; \theta')| \geqslant |F'(\theta'; \theta) - F'(\theta'; \theta')|, \forall \theta, \theta'.$$

If a unique cutoff equilibrium exists for all games (m, n) where $m + n = c < 1$ for a population with belief profile F, then it also exists for a population with belief profile F'.[4]

The numerical example below demonstrates the contents of these two corollaries.

Example 5. Consider game $(m, n) = (0.25, 0.5)$, whose payoff matrix is shown below.

	a_L	a_H
a_L	$0.5, 0.5$	$1 - \theta, \theta'$
a_H	$\theta, 1 - \theta'$	$0.25, 0.25$

Let the belief profile be

$$F(\theta'; \theta) = \begin{cases} 0, & 0 \leqslant \theta' < \frac{5\theta}{6}; \\ 6\theta' - 5\theta, & \frac{5\theta}{6} \leqslant \theta' \leqslant \frac{5\theta+1}{6}; \\ 1, & \frac{5\theta+1}{6} < \theta' \leqslant 1. \end{cases}$$

In words, a type-θ player believes that the population is uniformly distributed on the interval $[\frac{5\theta}{6}, \frac{5\theta+1}{6}]$ of length 1/6. We show in the following that this population is too self-similar for a cutoff equilibrium to exist.

First of all, note that for a player whose type $\theta < 0.5$, a_L is the dominant strategy. Similarly, for any player whose type $\theta > 0.75$, a_H is the dominant strategy. The candidate cutoff type

[4]The formal definition of one population being more self-similar than another is on page 28. It is stronger than $|F(\theta'; \theta) - F(\theta'; \theta')| \geqslant |F'(\theta'; \theta) - F'(\theta'; \theta')|$.

$\theta^* \in (0.5, 0.75)$ is indifferent between a_L and a_H:

$$F(\theta^*;\theta^*)0.5 + (1 - F(\theta^*;\theta^*))(1-\theta^*) = F(\theta^*;\theta^*)\theta^* + (1 - F(\theta^*;\theta^*))0.25$$

so that $\theta^* = 0.6$ is the only candidate for a cutoff type.

Since

$$F(0.6;\theta) = \begin{cases} 1, & \theta < 0.52; \\ 3.6 - 5\theta, & 0.52 \leqslant \theta \leqslant 0.72; \\ 0, & \theta > 0.72. \end{cases}$$

It is easy to check that for any $\theta \in (0.6, 0.75)$, action a_H is strictly inferior for these players. For example, take $\theta = 0.7$. $F(0.6; 0.7) = 0.1$. A player whose type $\theta = 0.7$ conjectures that her opponent plays the high action with probability 0.9, which is much higher than the conjecture a cutoff type holds. Therefore,

$$EU(a_L) = 0.1 \cdot 0.5 + 0.9 \cdot 0.3 = 0.32 > 0.295 = 0.1 \cdot 0.7 + 0.9 \cdot 0.25 = EU(a_H).$$

The strategy where $\theta^* = 0.6$ is the cutoff type does not constitute a cutoff equilibrium. Since it is the only candidate, a cutoff equilibrium does not exist for this game.

Suppose the population is less self-similar:

$$G(\theta';\theta) = \begin{cases} 0, & 0 \leqslant \theta' < \frac{\theta}{2}; \\ 2\theta' - \theta, & \frac{\theta}{2} \leqslant \theta' \leqslant \frac{\theta+1}{2}; \\ 1, & \frac{\theta+1}{2} < \theta' \leqslant 1. \end{cases}$$

Now, a type-θ player believes that the population is uniformly distributed on interval $[\frac{\theta}{2}, \frac{\theta+1}{2}]$. Similarly, for a player whose type $\theta < 0.5$, a_L is the dominant strategy; and for any player whose type $\theta > 0.75$, a_H is the dominant strategy. The candidate cutoff type $\theta^* \in (0.5, 0.75)$ is indifferent between a_L and a_H:

$$G(\theta^*;\theta^*)0.5 + (1 - G(\theta^*;\theta^*))(1-\theta^*) = G(\theta^*;\theta^*)\theta^* + (1 - G(\theta^*;\theta^*))0.25$$

So $\theta^* = 0.6$ is again the only candidate.

Moreover,

$$G(0.6;\theta) = \begin{cases} 1, & \theta < 0.2; \\ 1.2 - \theta, & 0.2 \leqslant \theta \leqslant 1. \end{cases} \implies \begin{cases} \theta \in (0.5, 0.6) \implies a_L \succ_\theta a_H \\ \theta \in (0.6, 0.75) \implies a_H \succ_\theta a_L \end{cases}$$

Therefore, $\theta^* = 0.6$ is indeed an equilibrium cutoff type.

Finally, I demonstrate that for even this less self-similar population, when the substitution effect becomes too strong, a cutoff equilibrium does not exist. Consider the following game:

	a_L	a_H
a_L	$-1, -1$	$1-\theta, \theta'$
a_H	$\theta, 1-\theta'$	$-2, -2$

The candidate cutoff type $\theta^* \in (0.5, 0.75)$ is indifferent between a_L and a_H:

$$G(\theta^*;\theta^*)0.5 + (1 - G(\theta^*;\theta^*))(1-\theta^*) = G(\theta^*;\theta^*)\theta^* + (1 - G(\theta^*;\theta^*))0.25$$

So again $\theta^* = 0.6$.

However, take a player of type $\theta = 0.7$. Since $G(0.6, 0.7) = 0.5$, this player conjectures that her opponent plays a_H with probability 0.5. With this conjecture,

$$EU(a_H) = -2 \cdot 0.5 + 0.7 \cdot 0.5 = -0.65 < -0.35 = 0.3 \cdot 0.5 + (-1) \cdot 0.5 = EU(a_L).$$

A cutoff equilibrium does not exist.

Another important observation from the previous example is that the candidate cutoff type is determined by players' self-perceived quantiles function $F(\theta;\theta)$. Suppose $F(\theta;\theta) = G(\theta;\theta)$ for all θ, as in this example; whenever a unique cutoff equilibrium exists for both populations, the cutoff type has to be the same. In other words, within the set of all games that are played with cutoff equilibrium, the implications of a model with self-similar beliefs are identical to those of a type-independent belief model where $B(\theta) = F(\theta;\theta)$ for all $\theta \in \Theta$. In order to identify whether a particular game is played with cutoff equilibrium, the analyst needs to observe a richer primitive.

4.3 Individual action profiles

An *individual action profile* $\{A_i\}_{i\in I}$ records how individual players behave in different games. This additional information enables analysts to do two things: First, instead of identifying the population distribution of types, analysts can identify each player's private type θ_i. Second, with an individual action profile, an analyst can identify which games are played by monotone pure strategies.

Definition 8. An individual action profile $\{A_i\}_{i\in I}$ is a collection of mappings. For each player $i \in I$, $A_i : \mathcal{G} \to \{0, 1\}$.

For example, $A_i(m, n) = 1$ denotes that in game (m, n) individual i plays the high action a_H.

Definition 9. We say that a type assignment $\{\theta_i\}_i$ is surjective if for any $\theta \in [0, 1]$ there exists $i \in I$ such that $\theta_i = \theta$.

Definition 10. We say that an individual action profile $\{A_i\}_i$ has a *self-similar representation* if there exist a surjective type assignment $\{\theta_i\}_i$ and a belief profile function $F : [0,1] \times [0,1] \to [0,1]$ such that

1. Fix any $\theta \in [0,1]$, $F(\cdot;\theta)$ is a c.d.f. on $[0,1]$.
2. F is continuous, consistent and self-similar.
3. For any game (m,n) and $\theta^* \in [0,1]$,

$$\theta_i > \theta^* \implies A_i(m,n) = 1, \quad \theta_i < \theta^* \implies A_i(m,n) = 0$$

if and only if

$$\theta > \theta^* \implies \theta F(\theta^*;\theta) + m(1 - F(\theta^*;\theta)) \geqslant nF(\theta^*;\theta) + (1-\theta)(1 - F(\theta^*;\theta))$$
$$\theta < \theta^* \implies \theta F(\theta^*;\theta) + m(1 - F(\theta^*;\theta)) \leqslant nF(\theta^*;\theta) + (1-\theta)(1 - F(\theta^*;\theta))$$

In words, we say that an individual action profile has a self-similar representation if we can find a belief profile F and type assignment $\{\theta_i\}_{i\in I}$ such that for any game that has cutoff equilibrium under belief profile F, play is in accordance with that equilibrium. I make two comments on this definition. First, θ^* under item 3 may not exist. Therefore, the definition is partial in the sense that it requires no connection between the model and the data for games that do not have a cutoff equilibrium. Second, this definition requires that whenever a cutoff equilibrium exists, it is played by the population. This admittedly strong assumption reflects equilibrium selection based on monotonicity, which was motivated by data in Section 4.2.

For a substantial subset of games, players' preferred actions do not depend on their conjectures about their opponents' actions. For these games, players' beliefs about their opponents play no role in determining the equilibria: The model predicts that when $m \geqslant 1$, players of all types find a_H to be their dominant strategy. Similarly, when $n \geqslant 1$, players of all types find a_L to be their dominant strategy. The intuition is exactly the same as before. For example, for entry games, when the revenue of both firms producing in the market is large enough, no matter how self-similar firms' beliefs are, all firms enter the market.

Property 1 (Dominance)**.** For all $i \in I$, $m \geqslant 1$ implies that $A_i(m,n) = 1$ and $n \geqslant 1$ implies that $A_i(m,n) = 0$.

Consider the following set of *belief-proof games*: $M^0 := \{(1-n, n) \mid n \in [0,1]\}$.

Table 13: Payoff matrix for a belief-proof game

	a_L	a_H
a_L	n, n	$1-\theta, \theta'$
a_H	$\theta, 1-\theta'$	$1-n, 1-n$

In such games each player's preference over the two actions depends only on her own payoff relevant type θ. Take any game $(1-n, n)$ with $n \in [0,1]$ and any player $i \in I$,

$$\theta_i > n \implies a_H \succ_i a_L, \quad \theta_i < n \implies a_L \succ_i a_H,$$

regardless of her conjecture over her opponent's action. Moreover, the cutoff equilibrium with $\theta^* = n$ as the cutoff type is the unique SBNE for game $(1-n, n)$. Therefore, when an individual action profile $\{A_i\}$ has a self-similar representation, then it must also satisfy the monotonicity property stated below.

Property 2 (Monotonicity)**.** For any player i, if $A_i(1-n, n) = 1$ for some $n \in [0,1]$, then $A_i(1-n', n') = 1$ for all $n' \leqslant n$.

Players' private types can be uniquely identified with belief-proof games, a very small subset of $\mathcal{G}$.

Proposition 6. *If an individual action profile $\{A_i\}_i$ is represented by $(\{\theta_i\}_i, F)$. Then*

$$\theta_i = \sup\{n | A_i(1-n, n) = 1\}, \ \forall i \in I.$$

Let $t_i \in [0,1] := \sup\{n \mid A_i(1-n, n) = 1\}$ denote the *revealed type* of player i. It is important to distinguish t_i from θ_i: t_i is a statistic derived from the primitive, while θ_i is a parameter in the model. Knowing players' revealed types, an analyst can identify in which games a cutoff equilibrium is played. Let set $M \subset \mathcal{G}$ be the set of *revealed* cutoff equilibria games:

$$M := \{(m, n) \in \mathcal{G} \mid t_i \leqslant t_j, A_i(m, n) = 1 \implies A_j(m, n) = 1.\}$$

By construction, $M^0 \subset M$.

Inside the set M, axioms *Convexity* and *Continuity* also carry over.

Property 3 (Convexity)**.** Take any two games $g, g' \in M \subset \mathcal{G}$. For any $\gamma \in [0,1]$, if $A_i(g) = A_i(g')$ for almost all $i \in I$ then $A_i(\gamma g + (1-\gamma)g') = A_i(g)$ for almost all $i \in I$.

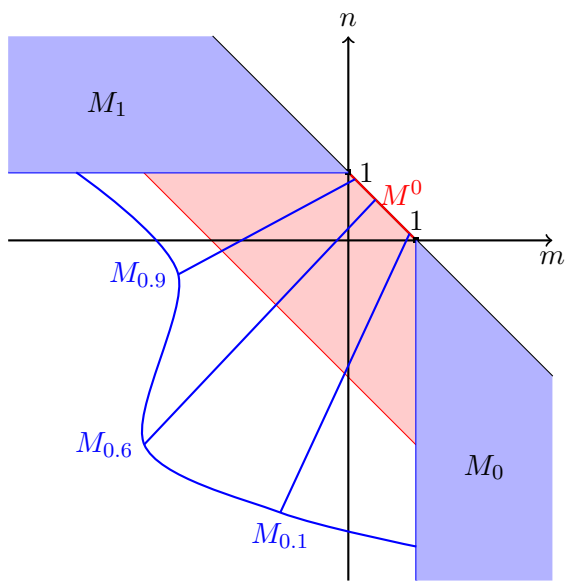
Property 4 (Closedness)**.** Take any sequence of games $\{g_k\}_{k \geqslant 1} \subset M \subset \mathcal{G}$ such that $g_k \to g \in \mathcal{G}$ as $k \to \infty$. If $A_i(g_k) = A_i(g_l)$ for all $k, l \geqslant 1$ and almost all $i \in I$, then $A_i(g) = A_i(g_1)$ for almost all $i \in I$. Therefore, $g \in M$.

Proposition 7. *Suppose an individual action profile $\{A_i\}_i$ is represented by $(\{\theta_i\}_i, F)$ then it has properties 1-4.*

Properties 1-4 are necessary conditions for an individual action profile to have a self-similar representation. In other words, an analyst can use these properties to test the hypothesis that an individual action profile has such a representation. Below, a typical set M that satisfies these properties is shown in the m-n plane. For any $t \in [0,1]$, $M_t \subset M$ denotes the set of games whose *revealed* cutoff type is t. Formally,

$$M_t := \{(m,n) \in \mathcal{G} \mid A_i(m,n) = A_i(1-t,t) \text{ for almost all } i \in I.\}$$

Figure 4: Partition of set M on m-n plain



This graph is very similar to the iso-value curves for the population action profiles when players hold type-independent beliefs. Both the iso-value curves and the M_t curves are straight lines and they fan out in the non-dominance region. Indeed, if players' private types are distributed according to H, then $A(g) = 1 - H(t)$ for all $g \in M_t$.

There are, however, subtle differences: First, the intersections of M_t curves with M^0 are evenly distributed on M^0. For example, $M_{0.9}$ intersects with M^0 at $(0.1, 0.9)$. This is not a coincidence as M_t directly records switching types. Moreover, players' revealed types are identified with their behavior in belief-proof games in M^0. Second, there is a boundary to the southwest of these curves. This is due to the fact that when players hold self-similar beliefs, a cutoff equilibrium does not exist when values of m and n are low. Properties above do not specify the shape of this boundary, but a self-similar representation implicitly does. This is one of the reasons why the above properties

do not qualify as sufficient conditions for the existence of a self-similar representation. However, suppose players' beliefs are only boundedly self-similar: Suppose there exists $N < \infty$, such that

$$\left| \frac{F(\theta'; \theta) - F(\theta'; \theta')}{\theta' - \theta} \right| \leqslant N, \quad \forall \theta, \theta' \in [0, 1].$$

Then there always exists a constant $c < 1$ such that

$$\{(m, n) \mid c \leqslant m + n \leqslant 1\} \subset M.$$

The shaded area above represents this band. From now on, I assume that such N exists.

To summarize, players' unobservable private types are uniquely identified through their behavior in belief-proof games. With these types, analysts can then identify which games are played with cutoff equilibria. Within this set of games (denoted by M), the behavioral implications of a self-similar belief model is exactly the same as that of a type-independent belief model. Moreover, the self-perceived quantile function $F(\theta; \theta)$ determines the cutoff types in these games exactly as the belief function $B(\theta)$ does in a type-independent belief model. Therefore, similar as for the type-independent belief model, the self-perceived quantile function $F(\theta; \theta)$ can be identified from observable play in games within M.

Proposition 8. *Suppose an individual action profile $\{A_i\}_i$ has full belief profile representations $(\{\theta_i\}, F)$ and $(\{\theta'_i\}, F')$, then $\theta_i = \theta'_i$ for all $i \in I$ and $F(\theta; \theta) = F'(\theta; \theta)$ for all $\theta \in [0, 1]$.*

Now that players have self-similar beliefs, one natural question is when we can say that one population is more self-similar than another. I propose the following definition.

Definition 11. We say that a belief profile F' is more self-similar than F if both F and F' are self-similar, and

1. For any $\theta, \theta' \in [0, 1]$, $|F'(\theta'; \theta) - F'(\theta'; \theta')| \geqslant |F(\theta'; \theta) - F(\theta'; \theta')|$; and
2. There exists $\theta^* \in (0, 1)$ such that

$$F'(\theta; \theta) \geqslant F(\theta; \theta), \forall \theta \leqslant \theta^*; \text{ and } F'(\theta; \theta) \leqslant F(\theta; \theta), \forall \theta \geqslant \theta^*.$$

The first requirement is straightforward: $|F(\theta'; \theta) - F(\theta'; \theta')|$ is a measure of how much beliefs differ when players are of different types. The larger the value, the more disagreement there is among the players. Since both F and F' are self-similar, it means that under belief profile F', the pattern that higher types believe there are more higher types becomes more significant. As established in Corollary 2, now that beliefs depend more on the type, cutoff equilibria cease to exist in more games.

The second requirement means that in the more self-similar population, players with low values of θ tend to overestimate their quantiles to a greater extent and those whose types are high tend to underestimate their quantiles to a greater extent. In the extreme case where players of all types believe they are the median, $F(\theta;\theta) = 1/2$ for all $0 < \theta < 1$. Therefore, when a game has low m and high n, it is the higher types who are close to being indifferent about a_H and a_L. Higher types underestimate their quantiles more severely in a more self-similar population. In other words, they believe there are more even higher types. Therefore, in a more self-similar population, more players would choose the low action in these games. The contrary happens in games with high m and low n. Lower types are close to being indifferent between a_H and a_L. Since they overestimate their quantile by a greater amount, they believe that there are even more lower types. More players would play a_H in these games.

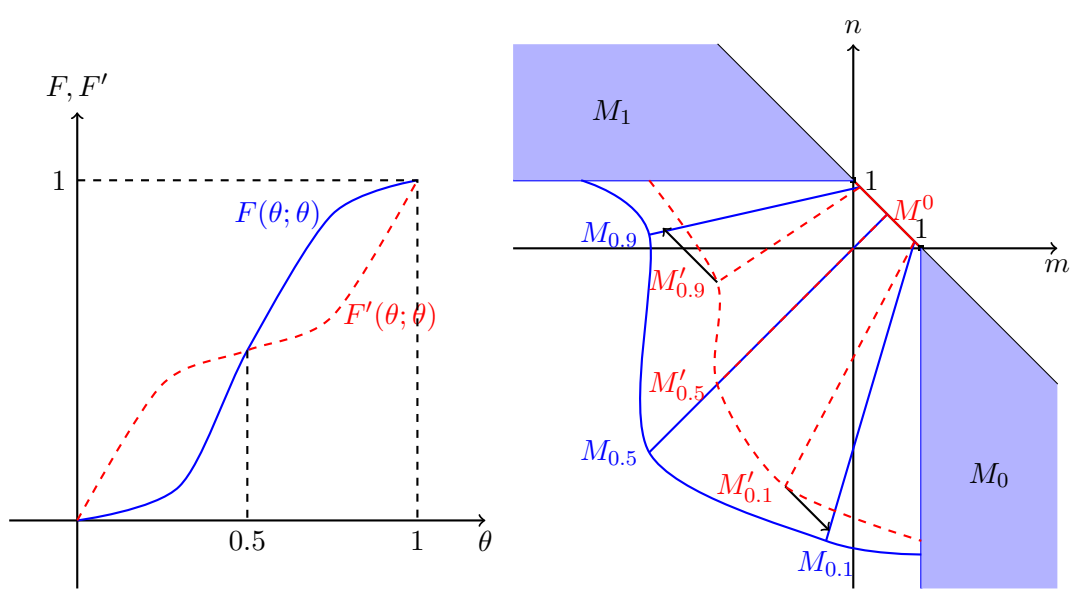
The following proposition formalizes the above intuitions.

Proposition 9. *Suppose individual action profiles $\{A_i\}_{i\in I}$ and $\{A`_j\}_{j\in J}$ have self-similar representations $(\{\theta_i\}, F)$ and $(\{\theta'_j\}, F')$ respectively. If belief profile F' is more self-similar than F, then for any $(m', n') \in M'_t$, there exists $(m, n) \in \mathcal{G}$ such that $m + n = m' + n'$ and $(m, n) \in M_t$. Moreover,*

$$t \leqslant \theta^* \implies m \leqslant m', n \geqslant n'; \text{ and } t \geqslant \theta^* \implies m \geqslant m', n \leqslant n'.$$

On the m-n plane, the M_t curves for a more self-similar population are less fanned out and the left-lower boundary of set M is pushed inwards.

Figure 5: Population F' is more self-similar than F



Empirically, more self-similar populations tend to herd more severely. For example, in teamwork games, if the incentive to work is very high, then self-similarity helps further increase the

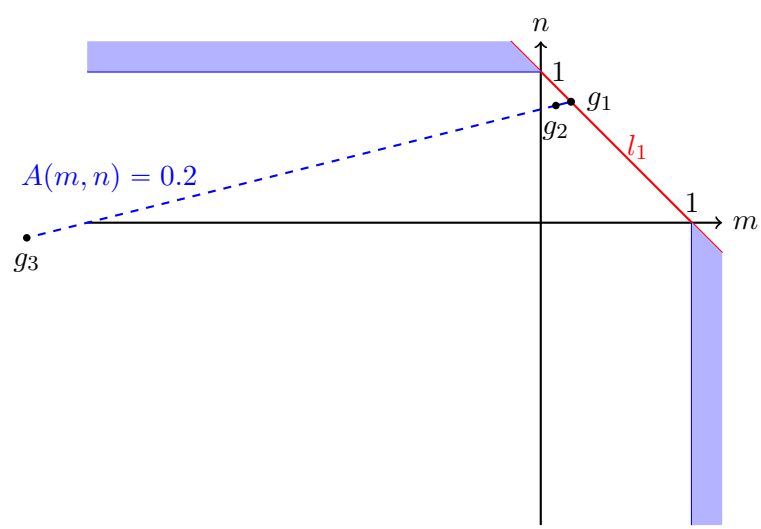
percentage of working individuals because players with high effort cost believe others are like them; conjecturing a high probability that her teammate would not work, she has a higher incentive to work. However, when the incentive to work is low, self-similarity further decreases the percentage of working individuals. Individuals with high effort cost would not work anyway because of this low incentive; those whose cost is low tend to think that their teammate also faces a low cost. They therefore overestimate the probability that their teammate will work, and are more likely to shirk.

4.4 Testing type-independent beliefs

As previously discussed, in games where values of m and n are low, cutoff equilibria cease to exist. In this subsection, I use a hypothetical example to illustrate that these games are useful for testing model specifications.

Suppose an analyst observes that $A(g_1) = A(g_2) = 0.2$ and wants to test the type-independent belief assumption. Recall the characterization result from Section 3: The iso-value curve where $A(m, n) = 0.2$ is the straight line that passes through these two points. In other words, the model implies that in all the games that lie on this line, 20% of the population plays a_H. The following example illustrates that, instead of games that lie in between g_1 and g_2, the analyst should look at game g_3.

Figure 6: Robustness check for the shared belief model



The payoff matrices of the three games are shown as below.

Figure 7: Payoff matrices of games g_1, g_2 and g_3

	a_L	a_H
a_L	$0.8, 0.8$	$0, 1$
a_H	$1, 0$	$0.2, 0.2$

(a) Game g_1

	a_L	a_H
a_L	$0.775, 0.775$	$0, 1$
a_H	$1, 0$	$0.1, 0.1$

(b) Game g_2

	a_L	a_H
a_L	$-1, -1$	$0, 1$
a_H	$1, 0$	$-7, -7$

(c) Game g_3

Suppose in population I, players' types are uniformly distributed on $[0, 1]$. It is easy to check that if all players hold the correct belief $U[0, 1]$; then as in all three games, $\theta^* = 0.8$ is the cutoff type. Therefore, 20% of the population plays a_H.

However, suppose that if $\theta_i \leqslant 0.9$, then player $i \in I$ holds the correct belief that her opponent's type is a random draw from the $[0, 1]$ interval. However, if $\theta_j > 0.9$, then player j believes that her opponent's type is a random draw from the $[0.8, 1]$ interval. Interpreting θ_i as i's degree of altruism as in Example 3 in Section 2, "selfish" players believe that their opponents' types are at least 0.8. Therefore, any player j whose type $\theta_j > 0.9$ believes that her opponent plays a_H with probability 1.

In game g_2, even this extreme belief does not eliminate the cutoff equilibrium. Take any player i those whose type $\theta_i \leqslant 0.9$, her belief about her opponent's action is $0.8a_L + 0.2a_H$. The analysis remains the same as when all players hold the correct belief: $\theta_i < 0.8$ implies that a_L is preferable, and $0.8 < \theta_i < 0.9$ implies that a_H is preferable. For those whose type $\theta_j > 0.9$, playing a_H generates utility of 0.1 and playing a_L generates $1 - \theta_j < 0.1$. In other words, since the payoff of outcome (a_H, a_H) is high enough, even with such extreme belief, the cutoff equilibrium remains. The same rationale works for all games that lie in between g_1 and g_2.

Table 14: Payoff matrix of game g_2

	a_L	a_H
a_L	$0.775, 0.775$	$0, 1$
a_H	$1, 0$	$0.1, 0.1$

Table 15: Utility of a type-θ row player game g_2

	a_L	a_H
a_L	0.775	$1 - \theta$
a_H	θ	0.1

However, the same amount of self-similarity in players' beliefs rules out a cutoff equilibrium in game g_3, whose payoff matrix is shown below. Recall that if $\theta_j > 0.9$, then player j believes that her opponent plays a_H with probability 1. Since $-7 < 1 - \theta_j$ for all $\theta_j \in (0.9, 1]$, player j finds a_L preferable.

Table 16: Payoff matrix of game g_3

	a_L	a_H
a_L	$-1, -1$	$0, 1$
a_H	$1, 0$	$-7, -7$

Table 17: Utility of a type-θ row player game g_3

	a_L	a_H
a_L	-1	$1 - \theta$
a_H	θ	-7

Moreover, I now demonstrate that in equilibrium, there cannot be 20% of the population who play the high action a_H in game g_3. First, observe that the equilibrium always takes the form of two-stage cutoff equilibrium. More specifically, let $\Theta_L = [0, 0.9]$ be the low type region and let $\Theta_H = (0.9, 1]$ be the high type region. Since players whose types are in the same region hold the same belief, my analysis in Section 3 implies that within each region, the equilibrium takes the form of a cutoff strategy.

In other words, there exists a cutoff type $\theta_1 \in [0, 0.9]$ such that $\sigma_i = a_L$ for all $\theta_i \in (0, \theta_1)$ and $\sigma_j = a_H$ for all $\theta_j \in (\theta_1, 0.9)$. Similarly, there exists a cutoff type $\theta_2 \in [0.9, 1]$ such that $\sigma_i = a_L$ for all $\theta_i \in (0.9, \theta_2)$ and $\sigma_j = a_H$ for all $\theta_j \in (\theta_2, 1)$.

Figure 8: A two-stage cutoff equilibrium

Suppose, contrary to the claim, that 20% of the population play a_H. Since players with type in Θ_L hold the correct belief about the population type distribution, $\Pr_\theta(a_H) = 0.2$ for all $\theta \in [0, 0.9]$. The previous discussion establishes that if a player holds this belief, then she is indifferent between a_H and a_L if and only if her type is $\theta = 0.8$. Players whose types are lower than 0.8 find a_L preferable and those whose types are between 0.8 and 0.9 find a_H preferable. Since 20% of the population plays a_H, it has to be the case that all those whose types are higher than 0.9 also play a_H. It is demonstrated above that this is not an equilibrium, a contradiction.

It turns out that the only equilibrium of game g_3 is where $\theta_1 = 0.733$ and $\theta_2 = 1$: Those whose types are between 0.733 and 0.9 play a_H and other players play a_L. In total, 16.7% of the population play a_H.

To summarize, in this example the analyst (successfully) rejects the null hypothesis of type-independent beliefs if she looks at game g_3. However, she fails to reject the common belief assumption if she looks at games that lie between g_1 and g_2. In general, games with low payoffs for outcomes (a_H, a_H) and (a_L, a_L) magnify the effect of possible self-similarity in beliefs and thus serve as better test games for the common belief assumption.

Remark

Following the same intuition, games with low m and n can also be used to distinguish self-similar beliefs about types from self-similar beliefs about actions, when players' conjectures can be observed.

Recall the positive correlation between actions and beliefs observed by Rubinstein and Salant (2016) and Charness et al. (2016). They ascribe this observation to players' *self-similar* beliefs about actions: Players believe that their opponents' actions are similar to their own even in contexts

where strategies are substitutes. In this paper, the definition of self-similarity is different: Players believe that their opponents' type, in this context, their degree of altruism, is similar to their own. In equilibrium, players' beliefs about their opponents' actions are products of their beliefs about others' types and the equilibrium action profile. The positive alignment between players' actions and beliefs is present only in monotone equilibria. In other words, my model predicts that this positive correlation disappears when m and n are very small. Interestingly, the "Battle of the Sexes" game can fall into this category.

Figure 9: Battle of the Sexes

	Opera	Football
Opera	$1, 0$	$-5, -5$
Football	$-5, -5$	$0, 1$

(a) Normalized Battle of the Sexes

	a_L	a_H
a_L	$-5, -5$	$0, 1$
a_H	$1, 0$	$-5, -5$

(b) Battle of the Sexes with actions relabeled

Note that for the row player (Pat), "Football" is labeled a_L and "Opera" is labeled a_H. At the same time, for the column player (Chris), "Opera" is labeled a_L and "Football" is labeled a_H. For Pat, who prefers "Opera" over "Football", "Football" is the action of being considerate by trying to coordinate on Chris's preferred activity. On the other hand, "Opera" is the more selfish action of trying to coordinate on Pat's own preferred activity. The opposite is true for Chris.

In a Battle of the Sexes game where both players strongly prefer to go to the same activity, if Pat chooses to go to "Football", then Pat probably conjectures that Chris also has a high chance of going to "Football". At first glance, this seems like a manifestation of self-similarity in actions. But once we introduce strategically appropriate symmetric labels, the opposite is true. By playing "Football" and believing that Chris would also play "Football", Pat is acting considerately and believing Chris to play selfish. Therefore, it would be interesting to observe both how people behave and what they believe when the Battle of the Sexes is framed in different ways.

5 Conclusion

This paper establishes that a population's aggregate behavior in games with strategic substitutes can be explained by a type-independent belief model if and only if it satisfies the axioms Dominance, Monotonicity, Continuity and Convexity. Self-similar beliefs about opponents' types, additionally, rationalize the positive correlation between players' actions and beliefs about their opponents' types. When players' beliefs about their opponents' actions are not observable, the behavioral implications of a self-similar belief model are exactly the same as those of a type-independent belief model for games where cutoff equilibria are played. An analyst can identify whether games are played with cutoff equilibria by keeping track of how each individual player behaves across games.

A General payoff structures

In this section, I discuss how the intuition carries to more general payoff structures. Players' payoffs are only assumed to satisfy the three key properties: first, higher types prefer the high action to a greater degree; second, the opponent's high action makes the high action less attractive; and third, players' payoffs are continuous with respect to their types.

More specifically, for any game g, let $u_g : X \times X \times \Theta \to \mathbb{R}$ denote a player's payoff given her private type, her own action and her opponent's action in game g, where $X := \{a_L, a_H\}$ denotes the set of actions, and the type space $\Theta := [\underline{\theta}, \bar{\theta}]$ is a compact subset of $\mathbb{R}$. For example, $u_g(a_L, a_H, \theta)$ is the payoff for a type-θ player to play a_L given that her opponent plays a_H. Players hold type-independent beliefs B about their opponents' types. The true type distribution is denoted by H.

In terms of observables, for each game g, the payoff function u_g is common knowledge for both players and is observable to the analyst, while players' private types are not. Suppose the analyst observes how a population behaves in a collection ($\mathcal{A}$) of symmetric 2×2 games.

For each game $g \in \mathcal{A}$, three key properties of the linear payoff structure are abstracted and assumed for its payoff function u_g. First, increasing differences in actions and types: Holding the opponent's action constant, the payoff gain of playing the high action is greater for higher types.

Definition 12. A payoff function u satisfies *increasing differences* in (a, θ) if for any $\theta < \theta'$ and $a \in \{a_H, a_L\}$,

$$u(a_H, a, \theta) - u(a_L, a, \theta) < u(a_H, a, \theta') - u(a_L, a, \theta').$$

Second, single crossing property of incremental returns in $(a, -a')$: If any type-θ individual finds the high action a_H preferable when her opponent plays high action a_H, she must also find the high action preferable when her opponent plays low action a_L. In other words, the opponent's high action discourages the high action of a player.

Definition 13. A payoff function u satisfies *single crossing property of incremental returns (SCP-IR) in* $(a, -a')$ if for any $\theta \in \Theta$,

$$u(a_H, a_H, \theta) - u(a_L, a_H, \theta) \geqslant (>) 0 \implies u(a_H, a_L, \theta) - u(a_L, a_L, \theta) \geqslant (>) 0$$

Third, for any game $g \in \mathcal{A}$, payoff function u_g is continuous.

Definition 14. We say that a payoff function $u : X \times X \times \Theta \to \mathbb{R}$ is continuous if $u(a, a', \theta)$ is continuous in θ for any $a, a' \in X$.

Observe that the games we studied before previous sections satisfy these conditions.

Claim 2. *Take any games $g \in \mathcal{G}$, its payoff function u_g is continuous and satisfies increasing differences in (a, θ) and SCP-IR in $(a, -a')$.*

When players hold type-independent beliefs and their payoff functions satisfy the above conditions, a unique SBNE exists and it is a cutoff equilibrium.

Claim 3. *For any belief B and any game $g \in \mathcal{A}$, a unique SBNE (σ, μ) exists. Moreover, the unique SBNE is a cutoff equilibrium.*

Definition 15. We say that a population action profile $A : \mathcal{A} \to [0,1]$ has a shared belief representation (H, B) if there exists continuous and strictly increasing c.d.f. functions $H, B : \Theta \to [0,1]$ such that take any game g, there exists a cutoff type $\theta^* \in \Theta$, such that $A(g) = 1 - H(\theta^*)$ if and only if

$$\begin{aligned} &B(\theta^*) \cdot u_g(a_H, a_L, \theta) + (1 - B(\theta^*)) \cdot u_g(a_H, a_H, \theta) \\ \geqslant (\leqslant) &B(\theta^*) \cdot u_g(a_H, a_L, \theta) + (1 - B(\theta^*)) \cdot u_g(a_H, a_H, \theta), \ \forall \theta > (<) \theta^*. \end{aligned}$$

The key intuition we get from previous sections is that for all games where the same fraction of the population plays the high action a_H, there exists a same cutoff type who is indifferent between a_H and a_L in all these games. Moreover, since the cutoff type player's belief about her opponent's type remains constant across games, her conjectures about her opponents' actions are also the same in these games. The following theorem states that given any payoff structure specification, a population action profile has a type-independent belief representation if such a type-belief pair can be found for each class of games where the same fraction of the population play the high action.

Theorem 4. *A population action profile A has a shared belief representation (H, B) if and only if $A(g) = 1(0) \iff u_g(a_H, a_H, \theta) \geqslant (\leqslant) u_g(a_L, a_H, \theta), \forall \theta \in \Theta$ and there exist a psudo-belief function $r : [0,1] \to [0,1]$ and a cutoff type function $t : [0,1] \to \Theta$ such that*

1. *Functions t and r are continuous and strictly decreasing with $r(0) = 1, r(1) = 0$ and $t(0) = \bar{\theta}, t(1) = \underline{\theta}$.*

2. *For any $\alpha \in (0,1)$, $A(g) = \alpha$ implies that*

$$\begin{aligned} &r(\alpha) \cdot u_g(a_H, a_L, \theta) + [1 - r(\alpha)] \cdot u_g(a_H, a_L, \theta) \\ = \ &r(\alpha) \cdot u_g(a_L, a_L, \theta) + [1 - r(\alpha)] \cdot u_g(a_L, a_L, \theta). \end{aligned}$$

As the name suggests, function t maps each fraction α to its corresponding cutoff type. In other words, α is the probability that a player has a type higher than $t(\alpha)$. Function r, on the other hand, maps each fraction α to the belief held by its corresponding cutoff type. An immediate corollary of Theorem 4 is stated as follows.

Corollary 3. *If function r and t are as stated in Theorem 4, then population action profile A is represented by (H, B), where $H(\theta) = 1 - t^{-1}(\theta)$ and $B(\theta) = r(t^{-1}(\theta))$.*

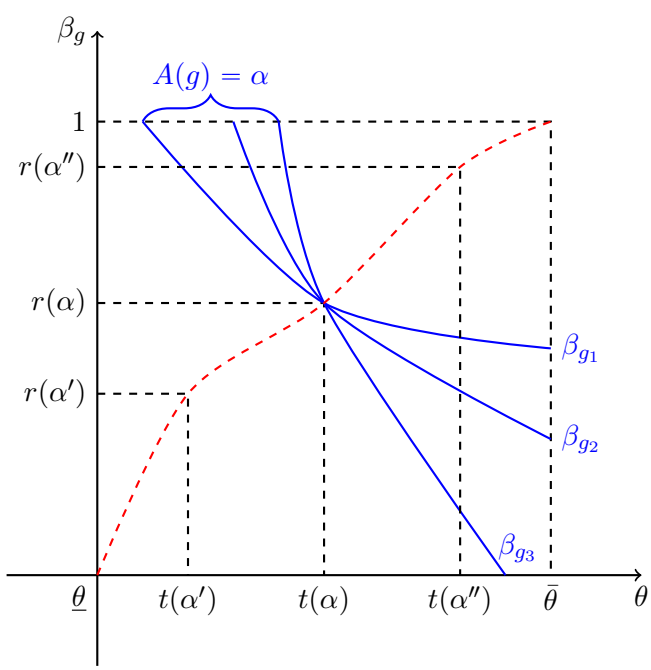
The contents of Theorem 4 and its corollary can be illustrated with the graph below. For any $g \in \mathcal{A}$ and $\theta \in \Theta$, let set

$$\beta_g(\theta) := \{\mu \in [0,1] \mid u_g(a_H, \mu a_L + (1-\mu)a_H, \theta) = u_g(a_L, \mu a_L + (1-\mu)a_H, \theta)\}$$

denote the set of conjectures for a type-θ player to be indifferent between a_H and a_L in game g. Pick any $\alpha \in (0,1)$; plot the indifference conjecture curves β_g for all games g such that $A(g) = \alpha$ on the θ-μ plane. The theorem states that all the indifference conjecture curves should intersect at a same point. The intuition is as follows: Suppose two games are played with the fraction of players playing a_H; their cutoff types must coincide. For the cutoff type, who is indifferent between a_H and a_L in both games, her belief about her opponent's action is also the same in both games, since the type-independent belief B is assumed to be constant over games. Therefore, all the indifference conjecture curves should pass the cutoff type and her conjecture. Let the intersection be $(t(\alpha), r(\alpha))$. $t(\alpha)$ is then the cutoff type, and $r(\alpha)$ the cutoff type's conjecture.

Suppose that $A(g') = \alpha' > \alpha = A(g)$. In other words, an analyst observes that more players play a_H in game g' than in game g. Then the cutoff type for game g' must be lower than that for game g. Players' shared belief about how many in the population plays a_L also becomes lower since the cutoff type is lower and belief B remains unchanged. Therefore, the intersection point $(t(\alpha'), r(\alpha'))$ is located to the southwest of $(t(\alpha), r(\alpha))$.

Figure 10: Illustration of Theorem 4



Example 6. Suppose an analyst studies individual players' behavior in games with monetary

payoffs and assumes that players have *charity preferences* (Charness & Rabin, 2002):

$$U_i(\pi_i, \pi_j) = \begin{cases} \theta \cdot \pi_i + (1-\theta) \cdot \pi_j & \pi_i > \pi_j; \\ \pi_i & \pi_i \leqslant \pi_j. \end{cases}$$

In words, a type-θ player assigns weight $1-\theta$ to her opponent's payoff only when her payoff is higher than her opponent's. When her own payoff is lower than her opponent's, she cares only about her own payoff.

Let the domain of games be $\mathcal{A} = \{(m,n) \mid m < 0, n < 1\}$. A typical element (m,n) denotes a 2×2 game whose monetary payoff and actual payoff matrices (when players have charity preferences) are shown as below.

(a) Monetary payoff matrix

	a_L	a_H
a_L	n, n	$0, 1$
a_H	$1, 0$	m, m

(b) Actual payoffs payoff matrix

	a_L	a_H
a_L	n, n	$0, \theta'$
a_H	$\theta, 0$	m, m

Then by definition,

$$\beta_g(\theta) = \begin{cases} \left\{\frac{-m}{\theta-n-m}\right\}, & \max\{n, 0\} \leqslant \theta \leqslant 1; \\ \varnothing, & \text{otherwise.} \end{cases}$$

Pick any $\alpha \in (0,1)$; the analyst can plot the β_g curves for all games such that $A(g) = \alpha$. For example, the analyst picks $\alpha = 0.8$. She observes that two games, g_1 and g_2, satisfy $A(g_1) = A(g_2) = 0.8$. In other words, she observes that in games $g_1 = (-0.6, 0.1)$ and $g_2 = (-0.3, 0.3)$, 80% of the population plays a_H. The payoff matrices of these two games are shown below.

(a) Payoff matrix for game g_1

	a_L	a_H
a_L	$0.1, 0.1$	$0, \theta'$
a_H	$\theta, 0$	$-0.6, -0.6$

(b) Payoff matrix for game g_2

	a_L	a_H
a_L	$0.3, 0.3$	$0, \theta'$
a_H	$\theta, 0$	$-0.3, -0.3$

When plotted, the indifference belief curves of g_1 and g_2 intersect at $(0.5, 0.6)$. In other words, a player with type $\theta = 0.5$ is indifferent between the two actions in both games if she conjectures that her opponent plays a_L with probability 0.6 and plays a_H with probability 0.4. The model then predicts that for any game g_3, if the same player ($\theta = 0.5$) is indifferent between the two actions under the same conjecture ($\mu(a_L) = 0.6, \mu(a_H) = 0.4$) in the new game g_3, then exactly 80% of the

population play a_H in g_3. Indeed, when $g_3 = (-0.1, -0.9)$, this is exactly the case. However, if the analyst observes that $A(g_3') = 0.8$ and $g_3' = (0.2, -0.8)$, then the data cannot be explained by a model where players have charity preferences and hold type-independent beliefs about each other.

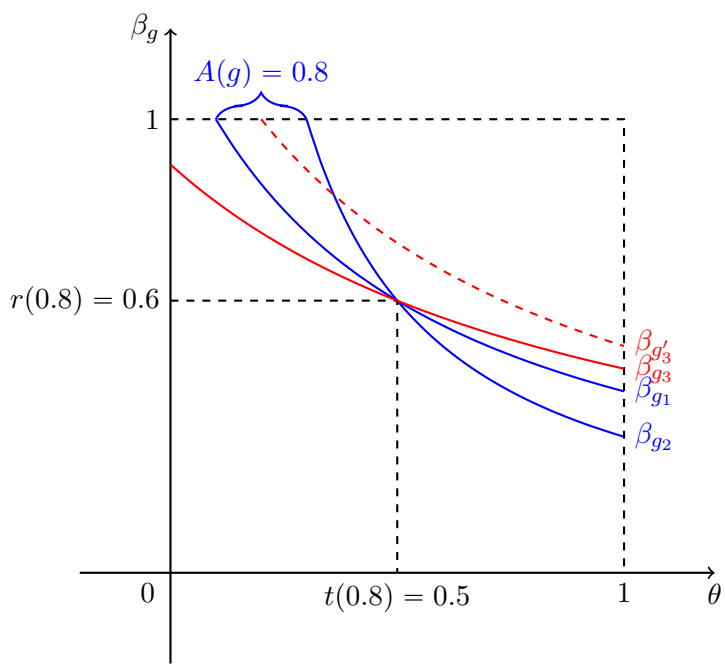
Intuitively, if we compare the payoffs of game g_3' with that of game g_1, the outcome of (a_H, a_H) generates a lower payoff while the outcome of (a_L, a_L) generates a higher payoff. We expect more players to play the low action. In other words, we expect that $A(g_3') < A(g_1)$. As is shown on the graph, in game g_3' player of type $\theta = 0.5$ is indifferent only if she holds a conjecture that her opponent plays a_L with probability higher than 0.6. At the conjecture of 0.6, she prefers a_L over a_H, which implies that cutoff type is higher than 0.5 and $A(g_3') > A(g_1)$.

(a) Payoff matrix for game g_3

	a_L	a_H
a_L	$-0.1, -0.1$	$0, \theta'$
a_H	$\theta, 0$	$-0.9, -0.9$

(b) Payoff matrix for game g_3'

	a_L	a_H
a_L	$0.2, 0.2$	$0, \theta'$
a_H	$\theta, 0$	$-0.8, -0.8$



B Comparison with magical thinking model

In this section, I compare the type-independent belief model with the magical thinking model proposed by Daley and Sadowski (2016). Specifically, I compare their behavioral implications in the context of lab experiments. Suppose subjects are presented with the following game:

Table 18: Lab experiment game

	$a_L(d)$	$a_H(c)$
$a_L(d)$	$0.1, 0.1$	$0, 1$
$a_H(c)$	$1, 0$	$0.2, 0.2$

The magical thinking model captures player's "illusion of control": Players believe that their action choice can influence their opponents' choice.(Daley & Sadowski, 2016) In the above game, by playing a_H, a player can get a higher payoff no matter what her opponent plays and at the same time increase the probability of outcome of (a_H, a_H) instead of (a_L, a_L). The magical thinking model predicts that all players play a_H (or c) in this game. Moreover, when m increases and n decreases, the statement remains true. This is a violation of the Monotonicity axiom.

When players hold altruistic preferences, there are always low types players who prefer the payoff pair $(0, 1)$ to $(0.2, 0.2)$, where the first item is their own payoff and the second item is their opponents' payoff. In other words, they are willing to sacrifice their own payoff to increase their opponents' payoff as long as the ratio is large enough. For example, a player of type $\theta = 1/2$ is willing to sacrifice her own payoff as long as her opponents' payoff increases by a larger amount. Therefore, the population is sensitive to the parameter change.

The fact that altruistic preference models violate Monotonicity axiom defined by Daley and Sadowski (2016) is established in their supplementary appendix. Therefore, the two models are behaviorally distinct.

C Proofs for Section 3

Proof of Proposition 1. First of all, applying the result of Theorem 3, a unique cutoff equilibrium exists. We only need to establish that any SBNE is a cutoff equilibrium.

Suppose $\sigma : [0, 1] \to \Delta\{a_L, a_H\}$ is a subjective Bayesian Nash equilibrium for game $g = (m, n)$. Suppose $\sigma_\theta(a_H) > 0$ for some $\theta \in [0, 1]$. Let $\mu \in [0, 1]$ be type-θ player's conjecture that her opponent plays a_H. Since all players have a shared belief, for any $\theta' > \theta$, type-θ' player's conjecture is also μ. Therefore,

$$\begin{aligned}
&\theta(1-\mu) + m\mu \geqslant n(1-\mu) + (1-\theta)\mu \\
\Longrightarrow\ &(1-m-n)(1-\mu) \geqslant 1-\theta-m \\
\Longrightarrow\ &(1-m-n)(1-\mu) > 1-\theta'-m \\
\Longrightarrow\ &\theta'(1-\mu) + m\mu > n(1-\mu) + (1-\theta')\mu
\end{aligned}$$

In words, a type-θ' player strictly prefers action a_H to a_L. Similarly, suppose $\sigma_\theta(a_L) > 0$, then $\sigma_{\theta'}(a_L) = 1$ for all $\theta' < \theta$. Therefore, any SBNE is a cutoff equilibrium. □

Before proving Theorem 1, I first establish that when a population action profile satisfies Axioms 1-4, then the following lemmas hold.

Lemma 1. *Fix any $c \leqslant 1$, for any $\alpha \in [0,1]$, there exists a unique $n \in [c-1,1]$ such that $A(c-n,n) = \alpha$.*

Proof. Fix any $c \leqslant 1$, it is immediate that function $q_c(n) := A(c-n,n) : [c-1,1] \to [0,1]$ is continuous and strictly decreasing. Moreover, $A(1,c-1) = 1$ and $A(c-1,1) = 0$.

Therefore, function q_c has a well-defined inverse function $q_c^{-1} : [0,1] \to [0,1]$, which is continuous and strictly decreasing. In other words, for any $\alpha \in [0,1]$, there exists a unique $n = q_c^{-1}(\alpha)$, such that

$$q_c(n) = A(c-n,n) = \alpha.$$

□

Lemma 2. *There exists functions $r, t : [0,1] \to [0,1]$ such that*

1. *Function t is continuous and strictly decreasing.*
2. *Function r is weakly decreasing.*
3. *For any $\alpha \in [0,1]$, $m < 1, n < 1$ and $m + n < 1$,*

$$A(m,n) = \alpha \implies r(\alpha) = \frac{1}{m+n-1} \cdot t(\alpha) + \frac{m-1}{m+n-1}.$$

Proof. Take $c = 1$ and apply the result of Lemma 1, function $q(n) := A(1-n,n) : [0,1]$ is continuous and strictly decreasing in n. Since $q(0) = 1$ and $q(1) = 0$, the inverse function $t := q^{-1} : [0,1] \to [0,1]$ is well defined. Moreover, function t is continuous and strictly decreasing in n.

Moreover, for any $\alpha \in [0,1]$, let

$$r(\alpha) := \frac{1}{m+n-1} t(\alpha) + \frac{m-1}{m+n-1},$$

where $A(m,n) = \alpha$, $m, n \leqslant 1$ and $m + n < 1$. First of all, want to show that value of $r(\alpha)$ does not depend on the choice of game (m,n).

Fix any $\alpha \in (0,1)$, let $(m,n) \neq (m',n')$ be such that $A(m,n) = A(m',n') = \alpha$. Then $m + n \neq m' + n'$, because otherwise monotonicity implies that $m = m'$ and $n = n'$. Therefore, without loss of generality assume that $m + n < m' + n' < 1$.

Let

$$r = \frac{1}{m+n-1} t(\alpha) + \frac{m-1}{m+n-1};$$
$$r' = \frac{1}{m'+n'-1} t(\alpha) + \frac{m'-1}{m'+n'-1}.$$

It implies that

$$(m+n-1)r = t(\alpha) + m - 1$$
$$(m'+n'-1)r = t(\alpha) + m' - 1$$

Since $m+n < m'+n' < 1$, there exists $\beta \in (0,1)$ such that

$$\beta(m+n) + (1-\beta)\cdot 1 = m'+n'.$$

Let

$$(m'', n'') := \beta(m,n) + (1-\beta)(1-t(\alpha), t(\alpha))$$

Then

$$\begin{aligned} n'' &= \beta \cdot n + (1-\beta)t(\alpha); \\ 1 - m'' &= \beta(1-m) + (1-\beta)t(\alpha); \\ m'' + n'' &= m' + n'; \end{aligned}$$

By convexity, $A(m'', n'') = \alpha$. Since $m'' + n'' = m' + n'$, it has to be the case that $m' = m''$ and $n' = n''$.

$$\begin{aligned} & n'' \cdot r + (1-m'')(1-r) \\ =& \big[\beta n + (1-\beta)t(\alpha)\big]r + \big[\beta(1-m) + (1-\beta)t(\alpha)\big](1-r) \\ =& \beta\big[n \cdot r + (1-m)(1-r)\big] + (1-\beta)\big[t(\alpha)\cdot r + t(\alpha)(1-r)\big] \\ =& t(\alpha) \\ \Longrightarrow & (m''+n''-1)r = t(\alpha) + m'' - 1. \end{aligned}$$

Compared with $(m'+n'+1)r = t(\alpha) + m' - 1$, the fact that $m'+n' = m''+n''$ and $m' = m''$ implies that $r = r'$. Therefore, function r is well-defined on $(0,1)$.

When $A(m,n) = \alpha = 0$, if $m, n \leqslant 1$ then the monotonicity of A implies that $n = 1$, which then implies that

$$r(0) = \frac{1}{m+n-1}t(0) + \frac{m-1}{m+n-1} = \frac{1}{m+1-1} + \frac{m-1}{m+1-1} = 1,$$

regardless of the choice of game (m,n). Similarly,

$$r(1) = \frac{1}{m'+n'-1}t(1) + \frac{m'-1}{m'+n'-1} = 0 + \frac{1-1}{1+n'-1} = 0,$$

where $A(m', n') = 1$, $m', n' \leqslant 1$ and $m' + n' < 1$. Therefore, function r is well-defined on $[0, 1]$.

Finally, we want to prove that function r is weakly decreasing. Take any $0 \leqslant \alpha' < \alpha \leqslant 1$, fix any $c < 1$, let n, n', l, l' be such that

$$A(c - n, n) = A(1 - l, l) = \alpha, \quad A(c - n', n') = A(1 - l', l') = \alpha'.$$

Then

$$\begin{aligned}
&l = t(\alpha), \quad l' = t(\alpha') \\
&(c-1)r(\alpha) = t(\alpha) + c - 1 - n = l + c - 1 - n \\
&(c-1)r(\alpha') = t(\alpha') + c - 1 - n' = l' + c - 1 - n'.
\end{aligned}$$

Suppose $r(\alpha') < r(\alpha)$, then

$$\begin{aligned}
&l' + c - 1 - n' > l + c - 1 - n \\
\Longrightarrow\ &l' - l > n' - n > 0.
\end{aligned}$$

Take any $\gamma > \frac{l'-l}{(l'-l)-(n'-n)} > 1$, let

$$x = \gamma n + (1 - \gamma)l, \text{ and } x' = \gamma n' + (1 - \gamma)l'.$$

Then $x > x'$.

Consider games

$$\begin{aligned}
g_1 &:= \gamma(c - n, n) + (1 - \gamma)(1 - l, l) = (c' - x, x), \text{ and} \\
g_2 &:= \gamma(c - n', n') + (1 - \gamma)(1 - l', l') = (c' - x', x');
\end{aligned}$$

where $c' := \gamma c + (1 - \gamma) \cdot 1 < c < 1$.

By monotonicity, $A(g_1) < A(g_2)$. However, convexity, continuity and monotonicity combined implies that $A(g_1) = \alpha$ and $A(g_2) = \alpha'$, as is demonstrated below. Take game g_1,

$$\begin{aligned}
&g_1 = (c' - x, x) = \gamma(c - n, n) + (1 - \gamma)(1 - l, l) \\
\Longrightarrow\ &(c - n, n) = \frac{1}{\gamma}(c' - x, x) + \left[1 - \frac{1}{\gamma}\right](1 - l, l)
\end{aligned}$$

Suppose $A(g_1) < \alpha$, then there exists $y < x$ such that $A(c' - y, y) = \alpha > A(g_1)$. It then implies

that

$$
\begin{aligned}
& A\left(\frac{1}{\gamma}(c'-y,y)+\left[1-\frac{1}{\gamma}\right](1-l,l)\right)=\alpha \\
\Longrightarrow\ & A\left(\frac{1}{\gamma}(c'-x,x)+\left[1-\frac{1}{\gamma}\right](1-l,l)\right)=A(c-n,n)<\alpha
\end{aligned}
$$

Contradiction is reached, since $A(c-n,n)=\alpha$. A similar contradiction can be reached if $A(g_1)>\alpha$.

By similar argument, $A(g_2)=\alpha'$. It contradicts the previously established statement that $A(g_1)<A(g_2)$. Therefore, it has to be the case that $r(\alpha')\geqslant r(\alpha)$. In other words, function r is weakly decreasing. Since $r(0)=1$ and $r(1)=0$, $r(\alpha)\in[0,1]$ for all $\alpha\in[0,1]$.

□

Lemma 3. *Functions r and t are continuous.*

Proof. Take any increasing sequence $\{\alpha_n\}$ such that $\alpha_n\to\alpha$. Sequences $t(\alpha_n)$ and $r(\alpha_n)$ are decreasing and lower bounded by 0. Therefore there exists $t,r\in[0,1]$ such that

$$t(\alpha_n)\to t\geqslant t(\alpha),\quad r(\alpha_n)\to r\geqslant r(\alpha)$$

Fix any $c<1$, for each α_n, there exists $l_n\in[c-1,1]$ such that $A(c-l_n,l_n)=\alpha_n$ and

$$l_n=t(\alpha_n)-(c-1)r(\alpha_n)+(c-1).$$

Moreover, l_n is decreasing and lower bounded by 0. Thus there exists $l\in[0,1]$ such that $l_n\to l$ as $n\to\infty$. By continuity of A, $A(c-l,l)=\lim_{n\to\infty}A(c-l_n,l_n)=\alpha$.

$$
\begin{aligned}
\lim_{n\to\infty} l_n &= \lim_{n\to\infty} t(\alpha_n)-(c-1)r(\alpha_n)+(c-1) \\
&= t+(1-c)r+(c-1) \\
&\geqslant t(\alpha)+(1-c)r(\alpha)+(c-1) \\
&= l.
\end{aligned}
$$

Therefore, $t=t(\alpha)$ and $r=r(\alpha)$.

Similarly, take any decreasing sequence β_n

$$
\begin{aligned}
& \lim_{n\to\infty}=\beta \\
\Longrightarrow\ & \lim_{n\to\infty} t(\beta_n)=t(\beta),\quad \lim_{n\to\infty} r(\beta_n)=r(\beta).
\end{aligned}
$$

Therefore, functions $r(\cdot)$ and $t(\cdot)$ are continuous. □

Proof of Theorem 1. **Necessity**: Suppose a population action profile has a shared belief representation (H, B).

Dominance: Consider game (m, n) where $m \geqslant 1$, then $m + n \leqslant 1$ implies that $n \leqslant 0$. For any $\theta, \mu \in [0, 1]$,

$$\begin{aligned} 1 - m - n \geqslant 0 \geqslant 1 - m - \theta &\implies (1 - m - n)\mu \geqslant 1 - m - \theta \\ &\implies \theta\mu + m(1 - \mu) \geqslant n\mu + (1 - \theta)(1 - \mu). \end{aligned}$$

The inequality is strict when $\theta > 0$. Therefore, for such games, $\theta^* = 0$ and $A(m, n) = 1 - H(\theta^*) = 1$. Similarly $A(m, n) = 0$ for all games with $n \geqslant 1$.

Monotonicity: Take any two games (m, n) and (m', n') such that $m < m' \leqslant 1$ and $n' < n \leqslant 1$. Let θ and θ' be their cutoff types. Then

$$\begin{aligned} &B(\theta) \cdot n + [1 - B(\theta)] \cdot (1 - \theta) = B(\theta) \cdot \theta + [1 - B(\theta)] \cdot m \\ \implies &B(\theta) \cdot n + [1 - B(\theta)] \cdot (1 - m) = \theta \end{aligned}$$

Similarly,

$$B(\theta') \cdot n' + [1 - B(\theta')] \cdot (1 - m') = \theta'$$

Suppose $\theta \leqslant \theta'$, then $B(\theta) \leqslant B(\theta')$ and

$$\begin{aligned} \theta &= B(\theta) \cdot n + [1 - B(\theta)] \cdot (1 - m) \\ &> B(\theta) \cdot n' + [1 - B(\theta)] \cdot (1 - m') \\ &\geqslant B(\theta') \cdot n' + [1 - B(\theta')] \cdot (1 - m') = \theta'. \end{aligned}$$

The second inequality is because $n' \leqslant 1 - m'$ and $B(\theta') \geqslant B(\theta)$. Contradiction is reached. Therefore, $\theta > \theta$ and $A(m, n) = 1 - H(\theta) < 1 - H(\theta') = A(m', n')$.

Continuity: Fix any $c \leqslant 1$, consider games $(c - n, n)$. When $c - 1 \leqslant n \leqslant 1$,

$$(1 - c)B(\theta_n) + \theta_n + c - 1 = n.$$

Let $\{n_k\}_{k \leqslant 1} \subset [c - 1, 1]$ be an increasing sequence that converges to $n \in (c - 1, 1]$. Let θ_{n_k} be the cutoff type for game $(c - n_k, n_k)$. It is shown that θ_{n_k} is also an increasing sequence. Since it

is bounded from above, let $\theta^* := \lim_{k\to\infty} \theta_{n_k}$. Then

$$\begin{aligned}
&(1-c)G(\theta_{n_k}) + \theta_{n_k} + c - 1 = n_k \\
\Longrightarrow\ &(1-c)G(\theta^*) + \theta^* + c - 1 = n = (1-c)G(\theta_n) + \theta_n + c - 1 \\
\Longrightarrow\ &\theta^* = \theta_n \\
\Longrightarrow\ &\lim_{k\to\infty} A(c-n_k, n_k) = \lim_{k\to\infty} 1 - H(\theta_{n_k}) = 1 - H(\theta^*) = 1 - H(\theta_n) = A(c-n, n)
\end{aligned}$$

The proof is similar for a decreasing sequence $\{n'_k\}_{k\to\infty}$ that converges to $n' \in [c-1, 1)$. Therefore, fix any $c \leqslant 1$, $A(c-n, n)$ is continuous on $[c-1, 1]$. Moreover, $A(c-n, n) = 1$ when $n \leqslant c-1$, ; and $A(c-n, n) = 0$ when $n \geqslant 1$. Therefore, $A(c-n, n)$ is continuous in n on $\mathbb{R}$.

Convexity: Suppose $A(m, n) = A(m', n') = \alpha$, then they have the same cutoff type θ^* where $\alpha = 1 - H(\theta^*)$. It is straightforward to check that θ^* is also the cutoff type for game $(\gamma m + (1-\gamma)m', \gamma n + (1-\gamma)n')$. Therefore,

$$A(\gamma m + (1-\gamma)m', \gamma n + (1-\gamma)n') = 1 - H(\theta^*) = \alpha, \forall \gamma \in [0, 1].$$

Sufficiency: It is established in the previous lemmas that function $t : [0, 1] \to [0, 1]$ is continuous and strictly decreasing with $t([0, 1]) = [0, 1]$. Its reverse function $t^{-1} : [0, 1] \to [0, 1]$ is thus well defined, continuous, and strictly decreasing. Let

$$H(\theta) := 1 - t^{-1}(\theta), \text{ and } B(\theta) := r\big(t^{-1}(\theta)\big), \forall \theta \in [0, 1].$$

The continuity of functions $t(\cdot)$ and $r(\cdot)$ imply that $H(\cdot)$ and $B(\cdot)$ are both continuous. Moreover, since function t is strictly decreasing, function H is strictly increasing. Since function r is weakly decreasing, function B is weakly increasing. Moreover,

$$\begin{aligned}
&H(0) = 1 - t^{-1}(0) = 0, \quad H(1) = 1 - t^{-1}(1) = 1; \\
&B(0) = r(t^{-1}(0)) = r(1) = 0, \quad B(1) = r(t^{-1}(1)) = r(0) = 1.
\end{aligned}$$

Pick any game (m, n) where $m + n < 1$, let $\theta^* = t(A(m, n))$. Then

$$\begin{aligned}
&B(\theta^*) = r(A(m, n)) = \frac{1}{m+n-1} t(A(m, n)) + \frac{m-1}{m+n-1} = \frac{\theta^* + m - 1}{m+n-1}; \\
&H(\theta^*) = 1 - A(m, n).
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\theta B(\theta^*) + m(1 - B(\theta^*)) \geqslant nB(\theta^*) + (1-\theta)(1 - B(\theta^*)), \forall \theta > \theta^*;\\
&\theta B(\theta^*) + m(1 - B(\theta^*)) \leqslant nB(\theta^*) + (1-\theta)(1 - B(\theta^*)), \forall \theta < \theta^*;\\
&A(m,n) = 1 - H(\theta^*).
\end{aligned}$$

For games of the shape $(1-n, n)$, let $\theta^* = n$. By definition, $\theta^* = t(A(1-n,n))$. The rest of the checking is trivial.

□

Proof of Proposition 2. Take any $t \in [0,1]$, consider game $(1-t, t)$, whether A has shared belief representation (H, B) or (H', B'), $\theta^* = t$. Therefore,

$$H(\theta^*) = 1 - A(1-t,t) = H'(\theta^*) \implies H(t) = H'(t).$$

Similarly, for $t \in [0,1]$, fix $c < 1$, there exists $n_c \in [0,1]$ such that $A(c - n_c, n_c) = A(1-t,t)$. Let θ_1 and θ_2 be the cutoff types for games $(c - n_c, n_c)$ and $(1-t,t)$ respectively. Then

$$1 - H(\theta_1) = A(c - n_c, n_c) = A(1-t,t) = 1 - H(\theta_2).$$

Since $H(\cdot)$ is strictly increasing, $\theta_1 = \theta_2 = t$. Since A has representations (H, B) and (H', B'), functions B and B' are both continuous and therefore

$$\begin{aligned}
&tB(t) + (c - n_c)(1 - B(t)) = n_c B(t) + (1-t)(1 - B(t))\\
&tB'(t) + (c - n_c)(1 - B'(t)) = n_c B'(t) + (1-t)(1 - B'(t))\\
\implies &B(t) = B'(t) = \frac{n_c + 1 - t - c}{1 - c}.
\end{aligned}$$

□

Proof of Theorem 3. Suppose for some game $(m,n) \in \mathcal{G}$, $A(m,n) > A'_H(m,n)$. Let θ^* and θ^{**} be the cutoff types for game (m,n) in the two populations. Then

$$\begin{aligned}
&1 - H(\theta^*) > 1 - H'(\theta^{**})\\
\implies &H'(\theta^*) \leqslant H(\theta^*) < H'(\theta^{**})\\
\implies &\theta^* < \theta^{**}
\end{aligned}$$

For all $\theta \in (\theta^*, \theta^{**})$,

$$(1 - m - n)B(\theta^*) \geqslant 1 - m - \theta \geqslant (1 - m - n)B'(\theta^{**})$$

If $m+n=1$, then $\theta = 1-m$ for all $\theta \in (\theta^*, \theta^{**})$. Contradiction. Otherwise,

$$\begin{aligned} & B(\theta^*) \geqslant B'(\theta^{**}) \geqslant B(\theta^{**}) \\ \Longrightarrow & \theta^* \geqslant \theta^{**} \end{aligned}$$

Contradiction is reached.

□

Proof of Proposition 4. **Necessity**: Suppose $H = H'$, $B \leqslant B'$ and A and A' are represented by (H, B) and (H', B') respectively. The fact that $A(m,n) \leqslant A'(m,n)$ for all (m,n) is proved in the previous result. For any game (m,n) where $m+n=1$, the cutoff type $\theta^* = n$ and $A(m,n) = 1 - H(n) = 1 - H'(n) = A'(m,n)$.

Sufficiency: Suppose $A(m,n) = A'(m,n)$ when $m+n=1$. Then for any $n \in [0,1]$, $A(1-n,n) = 1 - H(n) = A'(1-n,n) = 1 - H'(n)$, therefore $H(n) = H'(n)$ for all $n \in [0,1]$.

Pick any game $0 < \theta < 1$, let game (m,n) be such that $m+n<1$ and

$$A(m,n) = 1 - H(\theta).$$

Then $(1-m-n)B(\theta) = 1-m-\theta$. Suppose $B'(\theta) < G(\theta)$, then

On the other hand, $A'(m,n) \geqslant A(m,n)$ implies that $\theta \geqslant \theta'$, where θ' is the cutoff type for profile A'. By monotonicity, $\theta' \in (0,1)$ and

$$(1-m-n)B'(\theta') = 1-m-\theta'.$$

Therefore,

$$\begin{aligned} B'(\theta) \geqslant B'(\theta') &= \frac{1-m-\theta'}{1-m-n} \\ &\geqslant \frac{1-m-\theta}{1-m-n} = B(\theta), \forall 0 < \theta < 1 \end{aligned}$$

Finally, when $\theta = 0$ or 1, $B(\theta) = B'(\theta)$. □

Proof of Proposition 5. **Necessity**: $A'(m,n) \geqslant A(m,n)$ is proved previously. Suppose for two games $A(m,n) = A(m',n')$. Let θ and θ' be the cutoff types for the two games, then $\theta = \theta'$. Moreover, for all types $t < \theta = \theta'$,

$$nB(\theta) + (1-t)(1-B(\theta)) \geqslant tB(\theta) + m(1-B(\theta));$$

and for all $t > \theta = \theta'$,

$$nB(\theta) + (1-t)(1-B(\theta)) \leqslant tB(\theta) + m(1-B(\theta)).$$

Since $B(\theta) = B'(\theta)$ for all θ. The above inequalities also hold for function B'. It implies that if a population has B' as their shared belief, θ and θ' are also cutoff types for games (m,n) and (m',n'). Therefore,

$$A'(m,n) = 1 - H'(\theta) = 1 - H'(\theta') = A'(m',n').$$

Sufficiency:When $\theta = 0, 1$, $B(\theta) = B(\theta')$ is automatically true.

Pick any $\theta \in (0,1)$. Let (m,n) and (m',n') be two games where $m < 1, n < 1$ and

$$A(m,n) = A(m',n') = 1 - H(\theta) \in (0,1).$$

Then

$$\begin{aligned}
&(1-m-n)B(\theta) = 1-m-\theta, \quad (1-m'-n')B(\theta) = 1-m'-\theta \\
\Longrightarrow \theta &= \frac{n-n'}{(m+n)-(m'+n')}(1-m) + \frac{m-m'}{(m+n)-(m'+n')}n, \\
B(\theta) &= \frac{m-m'}{(m+n)-(m'+n')}.
\end{aligned}$$

Since $A'(m,n) = A'(m',n')$, $m < 1$ and $n < 1$, let θ' the shared cutoff types for these two games. Then

$$\begin{aligned}
&(1-m-n)B'(\theta') = 1-m-\theta', \quad (1-m'-n')B'(\theta') = 1-m'-\theta' \\
\Longrightarrow \theta' &= \frac{n-n'}{(m+n)-(m'+n')}(1-m) + \frac{m-m'}{(m+n)-(m'+n')}n, \\
B'(\theta') &= \frac{m-m'}{(m+n)-(m'+n')}.
\end{aligned}$$

It is obvious that $\theta = \theta'$ and it then further implies that $B'(\theta) = B(\theta)$. □

Proof of Theorem 2. **Necessity**: Suppose $A(m,n) = \alpha \in (0,1)$. Let θ^* be the cutoff type for game (m,n). Then $\theta^* \in (0,1)$ and

$$\begin{aligned}
&(1-m-n)B(\theta^*) + \theta^* = 1-m, \ B(\theta^*) = H(\theta^*) = 1-\alpha \\
\Longrightarrow &(1-m-n)(1-\alpha) + \theta^* = 1-m \\
\Longrightarrow &(1-m-n-\delta)(1-\alpha) + \theta^* = 1-m-\delta(1-\alpha) \\
\Longrightarrow &[1-(m+(1-\alpha)\delta)-(n+\alpha\delta)]B(\theta^*) + \theta^* = 1-(m+(1-\alpha)\delta).
\end{aligned}$$

Therefore, θ^* is also the cutoff type for game $(m + (1-\alpha)\delta, n + \alpha\delta)$ and

$$A(m + (1-\alpha)\delta, n + \alpha\delta) = 1 - H(\theta^*) = \alpha.$$

Sufficiency: Pick any $t \in (0,1)$ and $\delta < 0$, if $A(1-t,t) = \alpha$, then $H(t) = 1-\alpha$ since t is the cutoff type for game $(1-t,t)$. Moreover, let θ^* be the cutoff type for game $(1-t-(1-\alpha)\delta, t+\alpha\delta)$. Then

$$\begin{aligned}
&A(1-t-(1-\alpha)\delta, t+\alpha\delta) = \alpha \\
\implies &H(\theta^*) = 1-\alpha,\ (1-1-\delta)B(\theta^*) + \theta^* = 1-(1-t-(1-\alpha)\delta) \\
\implies &\theta^* = t,\ B(t) = 1-\alpha = H(t).
\end{aligned}$$

Since $H(0) = B(0) = 0$ and $H(1) = B(1) = 1$, $H(t) = B(t)$ for all $t \in [0,1]$.

□

D Proofs for Section 4

Proof of Claim 1. Suppose for some $\theta < \theta'$, $\alpha_\theta = a_H$ and $\alpha_{\theta'} = a_L$. Then $\mu_\theta > \mu_{\theta'}$ by assumption.

$$\begin{aligned}
&\mu_\theta \cdot u(a_H, a_H, \theta) + (1-\mu_\theta)u(a_H, a_L, \theta) \geqslant \mu_\theta \cdot u(a_L, a_H, \theta) + (1-\mu_\theta)u(a_L, a_L, \theta) \\
\implies &\mu_\theta \cdot \big[u(a_H, a_H, \theta) - u(a_L, a_H, \theta)\big] + (1-\mu_\theta)\big[u(a_H, a_L, \theta) - u(a_L, a_L, \theta)\big] \geqslant 0
\end{aligned}$$

The increasing difference property implies that

$$\mu_\theta \cdot \big[u(a_H, a_H, \theta') - u(a_L, a_H, \theta')\big] + (1-\mu_\theta)\big[u(a_H, a_L, \theta') - u(a_L, a_L, \theta')\big] > 0.$$

For the above expression to be positive, strategic substitution suggests that there are two cases.

1. $u(a_H, a_H, \theta') - u(a_L, a_H, \theta') \geqslant 0, u(a_H, a_L, \theta') - u(a_L, a_L, \theta') > 0;$
2. $u(a_H, a_H, \theta') - u(a_L, a_H, \theta') \leqslant 0, u(a_H, a_L, \theta') - u(a_L, a_L, \theta') > 0,.$

In either case, $\mu_\theta > \mu_{\theta'}$ implies that

$$\begin{aligned}
&\mu_{\theta'} \cdot \big[u(a_H, a_H, \theta') - u(a_L, a_H, \theta')\big] + (1-\mu_{\theta'})\big[u(a_H, a_L, \theta') - u(a_L, a_L, \theta')\big] > 0 \\
\implies &\mu_{\theta'} \cdot u(a_H, a_H, \theta') + (1-\mu_{\theta'})u(a_H, a_L, \theta') > \mu_{\theta'} \cdot u(a_L, a_H, \theta') + (1-\mu_{\theta'})u(a_L, a_L, \theta'),
\end{aligned}$$

which contradicts the fact that $\alpha_{\theta'} = a_L$.

□

Before proving Theorem 3, I first establish the necessary and sufficient condition for a cutoff equilibrium to exist for each game $(m, n) \in \mathcal{G}$.

Lemma 4. *For any game (m, n), if $m, n < 1$ and $m + n < 1$, then there exists $\theta^* \in (0, 1)$ such that*

$$(m + n - 1)F(\theta^*; \theta^*) = \theta^* + m - 1.$$

Moreover, a unique cutoff equilibrium exists for game (m, n) if and only if

$$\frac{F(\theta^*; \theta) - F(\theta^*; \theta^*)}{\theta - \theta^*} \geqslant \frac{1}{m + n - 1}, \forall \theta \neq \theta^* \in [0, 1].$$

Proof of Lemma 4. Consider equation

$$F(\theta; \theta) = \frac{\theta + m - 1}{m + n - 1}.$$

By assumption, the left hand side is continuous and increasing in θ. The right hand side is continuous and strictly decreasing. When $\theta = 0$, $RHS = \frac{m-1}{m+n-1} > 0 = LHS$. When $\theta = 1$, $RHS = \frac{m}{m+n-1} < 1 = LHS$. Therefore, there exists a unique $\theta^* \in (0, 1)$ such that

$$(m + n - 1)F(\theta^*; \theta^*) = \theta^* + m - 1.$$

Necessity: Suppose a unique cutoff equilibrium exists for game (m, n). Since $m, n < 1$, $\theta = 0$ and 1 cannot be the cutoff type. Let the cutoff type be $\theta' \in (0, 1)$. Then for all $\theta \in (\theta', 1] \neq \varnothing$,

$$F(\theta'; \theta)n + (1 - F(\theta'; \theta))(1 - \theta) \leqslant \theta F(\theta'; \theta) + m(1 - F(\theta'; \theta));$$

and for all $\theta \in [0, \theta') \neq \varnothing$,

$$F(\theta'; \theta)n + (1 - F(\theta'; \theta))(1 - \theta) \geqslant \theta F(\theta'; \theta) + m(1 - F(\theta'; \theta)).$$

By continuity,

$$F(\theta'; \theta)n + (1 - F(\theta'; \theta))(1 - \theta) \geqslant \theta F(\theta'; \theta) + m(1 - F(\theta'; \theta)) \implies \theta' = \theta^*.$$

Moreover, for all $\theta \in (\theta^*, 1]$,

$$\begin{aligned}
& F(\theta^*;\theta)n + (1 - F(\theta^*;\theta))(1-\theta) \leqslant \theta F(\theta^*;\theta) + m(1 - F(\theta^*;\theta)) \\
\Longrightarrow\ & (m+n-1)F(\theta^*;\theta) \leqslant \theta + m - 1 \\
\Longrightarrow\ & (m+n-1)\left[F(\theta^*;\theta) - F(\theta^*;\theta^*)\right] \leqslant \theta - \theta^* \\
\Longrightarrow\ & \frac{F(\theta^*;\theta) - F(\theta^*;\theta^*)}{\theta - \theta^*} \geqslant \frac{1}{m+n-1}.
\end{aligned}$$

Similarly, for all $\theta \in [0, \theta^*)$,

$$\begin{aligned}
& F(\theta^*;\theta)n + (1 - F(\theta^*;\theta))(1-\theta) \geqslant \theta F(\theta^*;\theta) + m(1 - F(\theta^*;\theta)) \\
\Longrightarrow\ & (m+n-1)F(\theta^*;\theta) \geqslant \theta + m - 1 \\
\Longrightarrow\ & (m+n-1)\left[F(\theta^*;\theta) - F(\theta^*;\theta^*)\right] \geqslant \theta - \theta^* \\
\Longrightarrow\ & \frac{F(\theta^*;\theta) - F(\theta^*;\theta^*)}{\theta - \theta^*} \geqslant \frac{1}{m+n-1}.
\end{aligned}$$

Sufficiency: Suppose for all $\theta \neq \theta^*$,

$$\frac{F(\theta^*;\theta) - F(\theta^*;\theta^*)}{\theta - \theta^*} \geqslant \frac{1}{m+n-1}.$$

Then for all $\theta > \theta^*$,

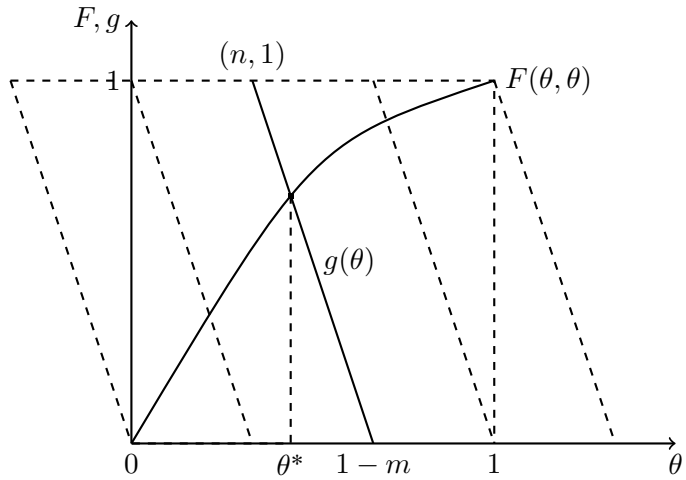
$$\begin{aligned}
& (m+n-1)\left[F(\theta^*;\theta) - F(\theta^*;\theta^*)\right] \leqslant \theta - \theta^* \\
\Longrightarrow\ & (m+n-1)F(\theta^*;\theta) \leqslant \theta + m - 1 \\
\Longrightarrow\ & F(\theta^*;\theta)n + (1 - F(\theta^*;\theta))(1-\theta) \leqslant \theta F(\theta^*;\theta) + m(1 - F(\theta^*;\theta)).
\end{aligned}$$

Therefore, all those players whose types are higher than θ^* find the high action a_H more preferable. Similarly, all those whose types are lower than θ^* prefer a_L. In other words, a cutoff equilibrium exists and its cutoff type is θ^*.

Let θ' be another cutoff type. The fact that $m, n < 1$ implies that $\theta' \neq 0, 1$. As is argued before $\theta' = \theta^*$. Therefore, the cutoff equilibrium is unique.

□

Figure 14: Proof sketch of Theorem 3



Proof of Theorem 3. **Necessity**: For any $\theta' \in (0, 1)$ and $c < 1$, there exists (m, n) such that $m + n = c, m < 1, n < 1$ and

$$(m + n - 1)F(\theta', \theta') = \theta' + m - 1.$$

Since a unique cutoff equilibrium exists for game (m, n), for any $\theta \neq \theta' \in [0, 1]$,

$$\frac{F(\theta'; \theta) - F(\theta'; \theta')}{\theta - \theta'} \geqslant \frac{1}{m + n - 1} = \frac{1}{c - 1}.$$

When $\theta' = 0$ or 1,

$$\frac{F(\theta'; \theta) - F(\theta'; \theta')}{\theta - \theta'} = 0 \geqslant \frac{1}{c - 1}.$$

Sufficiency: Pick any game (m, n) such that $m + n = c < 1$. If $m \geqslant 1$ or $n \geqslant 1$, a unique cutoff equilibrium exists. Otherwise, there exists $\theta^* \in (0, 1)$ such that

$$(m + n - 1)F(\theta^*; \theta^*) = \theta^* + m - 1.$$

Moreover,

$$\frac{F(\theta^*; \theta) - F(\theta^*; \theta^*)}{\theta - \theta'} \geqslant \frac{1}{c - 1} = \frac{1}{m + n - 1}.$$

A unique cutoff equilibrium exists for game (m, n). □

Proof of Corollary 1. Since a cutoff equilibrium exists for all games (m, n) where $m + n = c$, for any $\theta \neq \theta' \in [0, 1]$,

$$\frac{F(\theta'; \theta) - F(\theta'; \theta')}{\theta - \theta'} \geqslant \frac{1}{c - 1} > \frac{1}{c' - 1}.$$

It implies that cutoff equilibria exists for all games (m', n') where $m' + n' = c'$. $\square$

Proof of Corollary 2. Since cutoff equilibria exists for all c-games for population F, for all $\theta \neq \theta' \in [0, 1]$,

$$\begin{aligned}\frac{F'(\theta';\theta) - F'(\theta';\theta')}{\theta - \theta'} &= -\left|\frac{F'(\theta';\theta) - F'(\theta';\theta')}{\theta - \theta'}\right| \\ &\geqslant -\left|\frac{F(\theta';\theta) - F(\theta';\theta')}{\theta' - \theta}\right| = \frac{F(\theta';\theta) - F(\theta';\theta')}{\theta - \theta'} \geqslant \frac{1}{c-1}.\end{aligned}$$

Therefore, cutoff equilibria also exists for all c-games for population F'. $\square$

Proofs of Properties 1 and 2. Discussed in the text. $\square$

Proof of Proposition 6. Suppose $\{A_i\}_i$ is represented by $(\{\theta_i\}_i, F)$, then for any game $(1-n, n)$,

$$\theta_i > n \implies A_i(1-n, n) = 1, \text{ and } \theta_i < n \implies A_i(1-n, n) = 0.$$

Therefore, pick any $n \in [0, 1]$,

$$A_i(1-n, n) = 1 \implies \theta_i \geqslant n.$$

Thus, $\theta_i \geqslant t_i$.

Suppose $\theta_i > t_i$, then there exists $t \in (t_i, \theta_i)$ such that $A_i(1-t, t) = 0$. The fact that $A_i(1-t, t) = 0$ implies that $\theta_i \leqslant t < t_i$. Contradiction reached. $\theta_i = t_i$.

$\square$

Proof of Property 3. By definition, the fact that

$$\theta_i > t \implies A_i(m, n) = A_i(m', n') = 1; \quad \theta_i < t \implies A_i(m, n) = A_i(m', n') = 0.$$

implies that

$$\begin{aligned}\theta_i > t \implies & \theta_i F(t;\theta_i) + m(1 - F(t;\theta_i)) \geqslant nF(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)); \\ & \theta_i F(t;\theta_i) + m'(1 - F(t;\theta_i)) \geqslant n'F(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)); \\ \implies & \theta_i F(t;\theta_i) + m''(1 - F(t;\theta_i)) \geqslant n''F(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i));\end{aligned}$$

where $m'' = \alpha m + (1-\alpha)m'$ and $n'' = \alpha n + (1-\alpha)n'$ for some $\alpha \in [0, 1]$.

Similarly,

$$\theta_i < t \implies \theta_i F(t;\theta_i) + m''(1 - F(t;\theta_i)) \leqslant n''F(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)).$$

Since $(\{\theta_i\}_i, F)$ represents $\{A_i\}_i$,

$$\theta_i > t \implies A_i(m'', n'') = 1, \quad \theta_i < t \implies A_i(m'', n'') = 0.$$

□

Proof of Property 4. Suppose individual action profile $\{A_i\}_i$ has full belief profile representation $(\{\theta_i\}_i, F)$.

$$\theta_i > t \implies A_i(m_k, n_k) = 1; \quad \theta_i < t \implies A_i(m_k, n_k) = 0.$$

implies that

$$\begin{aligned}
\theta_i > t &\implies \theta_i F(t;\theta_i) + m_k(1 - F(t;\theta_i)) \geqslant n_k F(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)), \forall k. \\
&\implies \theta_i F(t;\theta_i) + m(1 - F(t;\theta_i)) \geqslant nF(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)), \forall k. \\
\theta_i < t &\implies \theta_i F(t;\theta_i) + m_k(1 - F(t;\theta_i)) \geqslant n_k F(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)), \forall k. \\
&\implies \theta_i F(t;\theta_i) + m(1 - F(t;\theta_i)) \leqslant nF(t;\theta_i) + (1-\theta_i)(1 - F(t;\theta_i)), \forall k.
\end{aligned}$$

Therefore,

$$\theta_i > t \implies A_i(m, n) = 1; \quad \theta_i < t \implies A_i(m, n) = 0.$$

□

Proof of Proposition 8. First of all, the lemma implies that $\theta_i = \theta_i' = t_i$ for all i.

Second, if for some (m, n),

$$\theta_i > t \implies A_i(m, n) = 1, \text{ and } \theta_i < t \implies A_i(m, n) = 0,$$

then

$$F(t;t) = \frac{m + t - 1}{m + n - 1} = F'(t;t).$$

□

Proof of Proposition 9. Pick any $(m', n') \in M_t'$, where $0 < t < 1$. Then

$$\begin{aligned}
&(1 - m' - n')F'(t;t) = 1 - m' - t; \text{ and} \\
&\left|\frac{F'(t;\theta) - F'(t;t)}{\theta - t}\right| \leqslant \frac{1}{1 - m - n}.
\end{aligned}$$

Let (m, n) be such that $m + n = m' + n'$ and

$$m = 1 - t - (1 - m' - n')F(t; t) \implies (1 - m - n)F(t; t) = 1 - m - t$$

Since belief profile F' is more self-similar than F,

$$\left|\frac{F(t;\theta) - F(t;t)}{\theta - t}\right| \leqslant \left|\frac{F'(t;\theta) - F'(t;t)}{\theta - t}\right| \leqslant \frac{1}{1 - m' - n'} = \frac{1}{1 - m - n}.$$

Therefore, t is also the cutoff type for game (m, n) with belief profile F. Since $\{A_i\}_{i \in I}$ is represented by $(\{\theta_i\}, F)$, $(m, n) \in M_t$.

Moreover, there exists $\theta^* \in (0, 1)$ such that, for all $t \leqslant \theta^*$, $F'(t; t) \geqslant F(t; t)$ and

$$m = 1 - t - (1 - m' - n')F(t; t) \leqslant 1 - t - (1 - m' - n')F'(t; t) = m'$$

Since $m + n = m' + n'$, $n \geqslant n'$. Similarly, when $t \geqslant \theta^*$, $m \geqslant m'$ and $n \leqslant n'$. □

E Proofs for Section A

Proof of Claim 2. For any game $g \in \mathcal{G}$,

$$\begin{aligned} u_g(a_H, a_H, \theta) - u_g(a_L, a_H, \theta) &= \theta + m - 1; \\ u_g(a_H, a_L, \theta) - u_g(a_L, a_L, \theta) &= \theta - n. \end{aligned}$$

Both differences are continuous and strictly increasing in θ.

Moreover, since $n \leqslant 1 - m$,

$$\begin{aligned} & u_g(a_H, a_H, \theta) - u_g(a_L, a_H, \theta) \geqslant (>) 0 \\ \implies & \theta + m - 1 \geqslant (>) 0 \\ \implies & \theta - n \geqslant (>) 0 \\ \implies & u_g(a_H, a_L, \theta) - u_g(a_L, a_L, \theta) \geqslant (>) 0. \end{aligned}$$

□

Before formally proving Claim 3, we first prove the following result:

Lemma 5. *Take any game $g \in \mathcal{A}$ and belief B, a unique cutoff equilibrium exists.*

Proof. Fix any game g, let

$$\delta(a_L, \theta) := u_g(a_H, a_L, \theta) - u_g(a_L, a_L, \theta) \text{ and } \delta(a_H, \theta) := u_g(a_H, a_H, \theta) - u_g(a_L, a_H, \theta)$$

denote the payoff gain switching from a_L to a_H when the opponent plays a_L and a_H respectively. Let

$$\theta_L := \operatorname*{argmin}_{\theta} \left|u_g(a_H, a_L, \theta) - u_g(a_L, a_L, \theta)\right| \text{ and } \theta_H := \operatorname*{argmin}_{\theta} \left|u_g(a_H, a_H, \theta) - u_g(a_L, a_H, \theta)\right|$$

denote the switching types when the opponent plays a_L and a_H respectively. Since payoff function u_g is continuous in θ and has increasing differences in (a, θ), the definition makes sense. Moreover, SCP-IR in $(a, -a')$ implies that $\theta_H \geqslant \theta_L$.

Observe that:

1. When $\theta < \theta_L$, $u_g(a_H, a_H, \theta) < u_g(a_L, a_H, \theta)$ and $u_g(a_H, a_L, \theta) < u_g(a_L, a_L, \theta)$, which implies that a_L is the strictly dominant strategy.

2. When $\theta_L < \theta < \theta_H$, $u_g(a_H, a_H, \theta) < u_g(a_L, a_H, \theta)$ and $u_g(a_H, a_L, \theta) > u_g(a_L, a_L, \theta)$ the player's preference over the two actions depends on her conjecture over her opponent's action.

3. When $\theta > \theta_H$, $u_g(a_H, a_H, \theta) > u_g(a_L, a_H, \theta)$ and $u_g(a_H, a_L, \theta) > u_g(a_L, a_L, \theta)$, which implies that a_H is the strictly dominant strategy.

Interval $I_n = [\theta_L, \theta_H]$ is thus the set of types who do not have a strictly dominant strategy. When $\theta_H = \theta_L$, $I_n = \{\theta_L\}$. Almost every type has a strictly dominant strategy. Moreover, $\theta^* = \theta_H = \theta_L$ is the unique cutoff type for a cutoff equilibrium.

When $\theta_H > \theta_L$, a player with type $\theta \in (\theta_L, \theta_H)$ weakly prefers a_H if and only if

$$\begin{aligned} &\mu \cdot u_g(a_H, a_L, \theta) + (1-\mu) \cdot u_g(a_H, a_H, \theta) \geqslant \mu \cdot u_g(a_L, a_L, \theta) + (1-\mu) \cdot u_g(a_L, a_H, \theta) \\ \iff &\mu \geqslant \frac{-\delta(a_H, \theta)}{\delta(a_L, \theta) - \delta(a_H, \theta)} := \gamma(\theta) \in (0, 1). \end{aligned}$$

where μ is the population's shared conjecture of a_L being played.

Consider equation

$$B(\theta) = \gamma(\theta).$$

Since

$$\gamma(\theta_L) \geqslant B(\theta_L), \quad \gamma(\theta_H) \leqslant B(\theta_H).^5$$

Since both functions are continuous, an intersection exists. Moreover, function B is increasing in θ and γ is strictly decreasing, the intersection is unique. Let the $\theta^* \in [\theta_L, \theta_H]$ be such that $B(\theta^*) = \gamma(\theta^*)$. It is easy to check that it is indeed a cutoff type.

It remains to show that it is also the unique cutoff type. Suppose there exists another cutoff type θ^{**}.

Note that when $\theta_L < \theta_H$, one of the following is true.

[5] If $\theta_L = \underline{\theta}$, then $\gamma(\theta_L) > 0 = F(\theta_L, \theta_L)$. If $\theta_L > \underline{\theta}$, then $\gamma(\theta_L) = 1 \geqslant F(\theta_L, \theta_L)$. Similarly for θ_H.

1. $\underline{\theta} < \theta_L < \theta_H < \bar{\theta}$;

2. $\underline{\theta} = \theta_L < \theta_H < \bar{\theta}$;

3. $\underline{\theta} < \theta_L < \theta_H = \bar{\theta}$;

4. $\underline{\theta} = \theta_L < \theta_H = \bar{\theta}$.

If $\theta^{**} = \underline{\theta}$, all types play a_H. This is only possible in Cases 2 and 4. In both cases, there exists $\epsilon > 0$, such that for all $\theta < \underline{\theta} + \epsilon$,

$$\gamma(\theta) = \frac{-\delta(a_H, \theta)}{\delta(a_L, \theta) - \delta(a_H, \theta))} > 0 = \mu.$$

This contradicts the fact that a_H is a best response.

Similarly, if $\theta^{**} = \bar{\theta}$, all types play a_L. This is only possible in Cases 3 and 4. In both cases, there exists $\epsilon' > 0$, such that for all $\theta > \bar{\theta} - \epsilon'$,

$$\gamma(\theta) = \frac{-\delta(a_H, \theta)}{\delta(a_L, \theta - \delta(a_H, \theta))} < 1 = \mu$$

This contradicts the fact that a_L is the best response.

Therefore, $\theta^* \in (\underline{\theta}, \bar{\theta})$, and it is a solution to equation $B(\theta) = \gamma(\theta)$. Therefore, $\theta^{**} = \theta^*$. The cutoff equilibrium is unique. □

Proof of Claim 3. It is established in the previous lemma that there always exists a unique cutoff equilibrium. It remains to show that every SBNE is a cutoff equilibrium. Suppose σ is an SBNE, take any $\theta < \theta' \in \Theta$,

$$\begin{aligned}
\sigma_\theta(a_H) > 0 &\implies \mu \cdot \delta(a_H, \theta) + (1-\mu)\delta(a_L, \theta) \geqslant 0 \\
&\implies \mu \cdot \delta(a_H, \theta') + (1-\mu)\delta(a_L, \theta') > 0 \\
&\implies \sigma_{\theta'}(a_H) = 1
\end{aligned}$$

Similarly, $\sigma_{\theta'}(a_L) > 0$ implies that $\sigma_\theta(a_L) = 0$. Therefore, every SBNE is a cutoff equilibrium. □

Proof of Theorem 4. **Necessity**: Suppose population action profile A is represented by (H, B). Let

$$t(\alpha) = H^{-1}(1-\alpha), \text{ and } r(\alpha) = B[t(\alpha)] = B[H^{-1}(1-\alpha)].$$

Since functions H and G are continuous and strictly increasing, functions t and r are continuous

and strictly decreasing. Moreover,

$$t(0) = H^{-1}(1) = \bar{\theta}, \quad t(1) = H^{-1}(0) = \underline{\theta}; \text{ and}$$
$$r(0) = B[t(0)] = B(\bar{\theta}) = 1, \quad r(1) = B[t(1)] = B(\underline{\theta}) = 1.$$

For any $\alpha \in (0,1)$, $A(g) = \alpha$ implies that $\theta^* = t(\alpha) = H^{-1}(1-\alpha) \in (\underline{\theta}, \bar{\theta})$ is the cutoff type for game g. Moreover, a type-θ player conjectures that with probability $B(\theta^*) = r(\alpha)$ that her opponent plays a_L. Therefore,

$$\begin{aligned}
&r(\alpha)u_g(a_H, a_L, \theta) + (1 - r(\alpha))u_g(a_H, a_H, \theta) \\
\geqslant &r(\alpha)u_g(a_L, a_L, \theta) + (1 - r(\alpha))u_g(a_L, a_H, \theta), \forall \theta \geqslant t(\alpha); \text{ and} \\
&r(\alpha)u_g(a_H, a_L, \theta) + (1 - r(\alpha))u_g(a_H, a_H, \theta) \\
\leqslant &r(\alpha)u_g(a_L, a_L, \theta) + (1 - r(\alpha))u_g(a_L, a_H, \theta), \forall \theta \leqslant t(\alpha).
\end{aligned}$$

By continuity of u_g, $r(\alpha) \in \beta_g(t(\alpha))$.

Finally, $A(g) = 1 \iff \theta^* = \underline{\theta}$. Every player in the population plays a_H and believes that her opponent plays a_H for sure. Therefore,

$$A(g) = 1 \iff u_g(a_H, a_H, \theta) \geqslant u_g(a_L, a_H, \theta), \forall \theta \in \Theta.$$

Similarly for $A(g) = 0$.

Sufficiency: Let $H(\theta) = 1 - t^{-1}(\theta)$ and $B(\theta) = r(t^{-1}(\theta))$. Since functions r and t are continuous and strictly decreasing, functions H and B are continuous and strictly increasing. It is also easy to check that $H(0) = B(0) = 0$ and $H(1) = B(1) = 1$.

Want to show that population action profile A is represented by (H, B). First of all, when $A(g) = 1$ or 0, the conditions are already assumed.

When $A(g) = \alpha \in (0,1)$, let $\theta^* = t(\alpha)$, then $\alpha = t^{-1}(\theta^*)$ and $B(\theta^*) = r(t^{-1}(\theta^*)) = r(\alpha) \in \beta_g(\theta^*)$. In other words,

$$\begin{aligned}
&B(\theta^*)u_g(a_H, a_L, \theta^*) + (1 - B(\theta^*))u_g(a_H, a_H, \theta^*) \\
=&B(\theta^*)u_g(a_L, a_L, \theta^*) + (1 - B(\theta^*))u_g(a_L, a_H, \theta^*) \\
\implies &B(\theta^*)[u_g(a_H, a_L, \theta^*) - u_g(a_L, a_L, \theta^*)] + (1 - B(\theta^*))[u_g(a_H, a_H, \theta^*) - u_g(a_L, a_H, \theta^*)].
\end{aligned}$$

For all $\theta > \theta^*$, by increasing difference property of u_g,

$$\begin{aligned}
&B(\theta^*)[u_g(a_H, a_L, \theta) - u_g(a_L, a_L, \theta)] + (1 - B(\theta^*))[u_g(a_H, a_H, \theta) - u_g(a_L, a_H, \theta)] > 0 \\
\implies &B(\theta^*)u_g(a_H, a_L, \theta^*) + (1 - B(\theta^*))u_g(a_H, a_H, \theta^*) \\
&> B(\theta^*)u_g(a_L, a_L, \theta^*) + (1 - B(\theta^*))u_g(a_L, a_H, \theta^*)
\end{aligned}$$

Similarly,

$$\begin{aligned}
\theta < \theta^* \implies &B(\theta^*)u_g(a_H, a_L, \theta^*) + (1 - B(\theta^*))u_g(a_H, a_H, \theta^*) \\
&< B(\theta^*)u_g(a_L, a_L, \theta^*) + (1 - B(\theta^*))u_g(a_L, a_H, \theta^*)
\end{aligned}$$

□

Proof of Corollary 3. Obvious from the above proof. □

References

Andreoni, J., & Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–753.

Athey, S. (2001). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, *69*(4), 861–889.

Bellemare, C., & Kröger, S. (2007). On representative social capital. *European Economic Review*, *51*(1), 183–202.

Bolle, F., & Kritikos, A. (1999). *Approaching fair behavior: self-centered inequality aversion versus reciprocity and altruism.* Diskussionspapier.

Charness, G., & Haruvy, E. (2002). Altruism, equity, and reciprocity in a gift-exchange experiment: an encompassing approach. *Games and Economic Behavior*, *40*(2), 203–231.

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*(3), 817–869.

Charness, G., Rigotti, L., & Rustichini, A. (2016). Social surplus determines cooperation rates in the one-shot prisoner's dilemma. *Games and Economic Behavior*, *100*, 113–124.

Costa-Gomes, M. A., Huck, S., & Weizsäcker, G. (2014). Beliefs and actions in the trust game: Creating instrumental variables to estimate the causal effect. *Games and Economic Behavior*, *88*, 298–309.

Costa-Gomes, M. A., & Weizsäcker, G. (2008). Stated beliefs and play in normal-form games. *The Review of Economic Studies*, *75*(3), 729–762.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and economic behavior*, *46*(2), 260–281.

Daley, B., & Sadowski, P. (2016). Magical thinking: A representation result. *Theoretical Economics, Forthcoming*.

Fehr, E., Fischbacher, U., Von Rosenbladt, B., Schupp, J., & Wagner, G. G. (2003). A nation-wide laboratory: examining trust and trustworthiness by integrating behavioral experiments into representative survey.

Frick, M., Iijima, R., & Ishii, Y. (2018). Dispersed behavior and perceptions in assortative societies.

Gale, D. (1950). Compact sets of functions and function rings. *Proceedings of the American Mathematical Society*, *1*(3), 303–308.

Glaeser, E. L., Laibson, D. I., Scheinkman, J. A., & Soutter, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, *115*(3), 811–846.

Marks, G., & Miller, N. (1987). Ten years of research on the false-consensus effect: An empirical and theoretical review. *Psychological bulletin*, *102*(1), 72.

McAdams, D. (2003). Isotone equilibrium in games of incomplete information. *Econometrica*, *71*(4), 1191–1214.

Milgrom, P. R., & Weber, R. J. (1985). Distributional strategies for games with incomplete information. *Mathematics of operations research*, *10*(4), 619–632.

Rubinstein, A., & Salant, Y. (2016). "isn't everyone like me?": On the presence of self-similarity in strategic interactions. *Judgment and Decision Making*, *11*(2), 168.

Sapienza, P., Toldra-Simats, A., & Zingales, L. (2013). Understanding trust. *The Economic Journal*, *123*(573), 1313–1332.

Stanfield, B., Greig, F., Herrmann, B., & Zeckhauser, R. (2008). Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *The American Economic Review*, *98*(1), 294–310.

Van Zandt, T., & Vives, X. (2007). Monotone equilibria in bayesian games of strategic complementarities. *Journal of Economic Theory*, *134*(1), 339–360.