A DYNAMIC FOUNDATION OF THE MAXMIN CRITERION

IN-KOO CHO AND AKIHIKO MATSUI

ABSTRACT. This paper rigorously examines a dynamic foundation of the maxmin criterion (also known as Rawlsian criterion), that selects a social outcome which maximizes the payoff of the worst off agent. In order to understand how the behavior of individual agents shapes up a social preference through repeated long term interactions among agents, we analyze a dynamic matching process in which a long term relationship is embedded. A society is populated with a finite number agents with two different roles, say row and column agents. Two agents are matched, and initiate the long term relationship if both agents agree to do so, which is then subject to a small probability of break-up to dump the agents back to the matching process.

We focus on the outcomes sustained by a threshold decision rule slightly perturbed to incorporate sympathy in the spirit of [8]. For a general class of two person games, as the probability of continuing the long term relationship converges to 1, all agents in the economy almost always play the strongly Pareto efficient Rawlsian outcome, the most egalitarian outcome among strongly Pareto efficient outcomes. The Rawlsian criterion is generated through repeated long term interactions among agents in a decentralized fashion.

KEYWORDS: Matching, Long-term relationship, Social norm, Procedural rationality, Sympathy, Rawlsian criterion

1. INTRODUCTION

This paper rigorously examines a dynamic foundation of the Rawlsian criterion, which selects a social outcome that maximizes the payoff of the worst-off agent. We build a canonical model of social interactions, which unambiguously predicts a social outcome over a broad class of games to understand how individual behaviors shape up a social preference, which is significantly different from the selfish preference that drives the individual behavior.

The Nash bargaining solution is a prominent example of the selection criterion of a social outcome. In addition to the axiomatic foundation by [14], the solution concept has a non-cooperative game theoretic foundation ([22] and [18]) and an evolutionary foundation ([25]). More importantly, the Nash bargaining solution is evidently sensible if the underlying social situation is symmetric. However, it is less obvious whether the same criterion identifies an equally sensible outcome in an asymmetric bargaining problem. Indeed, many
Experimental evidences indicate that the split of surplus is often made around the efficient egalitarian outcome even in a game in which two players are clearly asymmetric (e.g., [17]). Even in a dictator game where the first mover can dictate the entire surplus in case of disagreement, a simple cheap talk between the dictator and the other party often leads to an equal split of the surplus ([24]). We consider these observations as a rationale to investigate the social preference that selects the most egalitarian outcome in the Pareto frontier, which is essentially the Rawlsian criterion.

A social preference is represented by the social outcome. Our goal is to understand the foundation of a social preference which is represented by the maxmin outcome over a general class of games. We view a social outcome as a steady state of a repeated long term interactions among agents in the society. We shall construct a decentralized matching model, which induces the maxmin outcome over a general class of two person games as a steady state of the long run outcomes of a social dynamics.

We consider a society populated with two groups of agents, row and column agents. In each period, unmatched row agents and column agents are searching for a long term relationship, while other agents are already in the long term relationship. After an unmatched row agent is randomly matched to another unmatched column agent, a relation-specific pair of payoffs is realized. The two agents form a long term relationship if both agree upon the long term stream of payoff vectors. Otherwise, both agents return to their respective pools of unmatched agents after receiving the realized payoff. If the two agents agree to form a long term relationship, then each agent receives the agreed payoff in the current period, and in the next period, the same pair is formed and the same payoffs are realized, though it is subject to a random shock with a small probability. The long term relationship lasts until one of the agents terminates it. When the long term relationship is dissolved, both agents join the respective groups of unmatched agents.

Our approach to investigate a social preference differs from a typical method widely used in behavioral economics. A typical approach of behavioral economics modifies the underlying preference of the individual decision maker by incorporating the payoff or the perception of the other player, and identify a solution from the game with a modified payoff function to interpret the experimental data ([1]). While offering a flexible, yet parsimonious, platform to investigate the behavior of economic agents, the same approach has been subject to criticism for the liberal use of the free parameters, allowing substantial departure from the standard economic analysis.\(^1\)

Instead of directly encoding the social behavior into the preference of the individual player, we largely maintain the fundamental assumption that each individual is selfish. Actually, we perturb the intrinsic of the individual player, but do so in a very disciplined manner, ensuring that the “size” of perturbation is arbitrarily small. Even if a social outcome indicates the behavior that significantly differs from selfish behavior, we cannot, and should not, conclude that the underlying preferences of individual players differ significantly from the conventional selfish preferences. Our model illuminates the dynamic mechanism through which an arbitrarily small departure from a fully selfish behavior generates a social preference, under which each individual behaves as if he is concerned of the

---

\(^1\)See also [7], who stress the value of maintaining various limits on economic models.
welfare of the worst off agent. In this sense, a social preference is a reduced form of the steady state arising from social dynamics.

To describe the “small” departure from a selfish rational player, let us imagine an agent, who is not fully aware of the structure of the society, or does not have perfect foresight, which is essential in computing an equilibrium strategy. When a player is matched with someone else, the decision to start a long term relationship requires to solve an extremely complex fixed point problem. Instead, each player estimates the value function through the simple average of his past payoff. The past average payoff is a consistent estimator of the value function of the continuation play only in a steady state. Since the agent treats the simple average as a consistent estimator following every history, he is certainly boundedly rational. However, his behavior and his reasoning in a steady state are very much consistent with a rational behavior. A main thrust of the analysis is to show the existence of a unique steady state.

An agent does not have perfect foresight and consequently, has to estimate the value function. Yet, his decision rule is procedurally rational ([15]), as his decision to initiate a long term relationship is based upon the same reasoning process, which he would have followed in an equilibrium. If the society is in a steady state, the long run behavior of the agent is observationally equivalent to an equilibrium behavior, as his estimator of the value function converges to the true value function.

While using the past average of his own past payoffs as a proxy for the value function, each agent uses with a small probability, the past average of his partner’s payoffs. This perturbation of the decision rule motivated by sympathy in the spirit of [8], who define sympathy as communication of sentiments and passions. When an agent is matched with another agent from a different population, his decision to initiate the long term relationship is mostly based on his selfish interest, consistent with laws of nature ([8]). But, with a small probability, he follows the rules of good-breeding, “in order to prevent the opposition of men’s pride, and render conversation agreeable and inoffensive.” (Section 2, Part 3, Book III, [8]) The pair of potential partners communicate about their threshold levels. With a small probability, each agent independently uses the other agent’s estimated value function in deciding on the long term relationship.

Two important elements of the social dynamics are that the long term relationship is voluntary, and can be terminated at any time unilaterally, and that the process to form a long term relationship is decentralized. The social dynamics in our model has a number of important features found in [5] and [6], where a repeated game with an option to drop out is investigated. Our social dynamics differs significantly from a class of matching models based on [19], because a successful match can last for an extended periods in contrast to the immediate dissolution as in [19]. As a result, the number of players in the pool of unmatched players is endogenously determined and smaller than the total number of players, while in [19], every player in the society is always available for matching. The endogenous number of available players for matching poses major challenges, which do not appear in a matching model where the matched pair is immediately separated after the completion of trade [19].
Like [6] and [5], our model generates a large number of equilibrium outcomes. An agent in the long term relationship has to compare the benefit of following the agreed behavior with the gain from dissolving the relationship to search for another relationship. We can make the prospect of finding a new relationship as pessimistic as possible by manipulating the acceptance rule of each agent. By requiring each agent to accept only the long term relationship that generates an average payoff close to the status quo, a agent has to reject a more profitable long term relationship in an equilibrium. We focus on the set of equilibrium outcomes that can be sustained by the decision rule which satisfies the two properties described above: threshold rule, and sympathy.

We demonstrate that sympathy combined with the threshold rule leads to the most egalitarian outcome among Pareto efficient outcomes, which is precisely the efficient element among Rawlsian outcomes, which maximize the payoff of the worst off agent in the society. More precisely, as the probability of the continuing the long term relationship converges to 1, almost every agent in the society almost always behaves as if he chooses an action according to the Rawlsian criterion whenever he is matched to another agent. This convergence result holds for any two person game, which the long term relationship between a pair of agents is built upon.

As we investigate the dynamic foundations of a social preference, our project is related to [16], although our focus is on the social preference in contrast to (individual) preference investigated therein. As we model social dynamics as a matching process among boundedly rational agents, our exercise is related to social learning models (e.g., [9], [25]). Since a social preference can select a particular social outcome, our project offers a useful insight toward the dynamic foundations of social norms and conventions ([23], [3], [11], [21], [4], [20], and [2]).

The rest of the paper is organized as follows. Section 2 formally describes the basic model and investigate the properties of threshold equilibria. In section 3, we perturb the basic game by incorporating the sympathetic behavior and characterize the maxmin outcome as the equilibrium steady state of symmetric threshold equilibria. We demonstrate in section 4 that the maxmin outcome is the only stable point of the learning dynamics in which each player estimates the value function recursively, and chooses his action based on the estimated value function. Section 5 concludes the paper.

2. Basic Model

2.1. Environment. Time is discrete, 1, 2, . . . , and its generic element is written as t. Let \( I = I^r \cup I^c \) be the set of agents where \( I^r = \{1, 2, \ldots, n\} \) is the set of anonymous row agents, and \( I^c = \{n + 1, \ldots, 2n\} \) is the set of anonymous column agents. In each period, each row agent is matched with a column agent, and vice versa. There are two pools of single agents, one for row agents, and the other for column agents. The set of row (resp. column) agents in the beginning of period \( t \) is denoted by \( U_r^t \) (resp. \( U_c^t \)).

In the initial period, every player is in the corresponding pool of singles:

\[
U_r^1 = \{1, \ldots, n\} \quad \text{and} \quad U_c^1 = \{n + 1, \ldots, 2n\}.
\]
For \( t \geq 1 \), players in \( U^r_t \) are randomly matched with some player in \( U^c_t \). We write \( U_t = U^r_t \cup U^c_t \). We assume that \( \forall i \in U^r_t, \forall j \in U^c_t \),

\[
(2.1) \quad P(i \text{ meets } j) = \frac{1}{\#U^c_t} \quad \text{and} \quad P(j \text{ meets } i) = \frac{1}{\#U^r_t}
\]

where \( \#U \) is the number of players in set \( U \).

Define \( U_t = U^r_t \cup U^c_t \) as the set of unmatched players, or the pool of singles. The set \( I \setminus U_t \) consists of the agents who agree to stay with the same partner in the previous period. We say that these agents are the matched pairs.

Let us denote by \( \mathcal{P} \) a partition of \( I \setminus U \) into pairs of long term partners,

\[
\mathcal{P} = \{\{i_1,j_1\}, \ldots, \{i_k,j_k\}\},
\]

where we have \( i_s \in I^r \) and \( j_s \in I^c \) (\( s = 1, \ldots, k \)). Let

\[
q_t = (U^r_t, U^c_t, \mathcal{P}_t)
\]

be a coalitional structure at time \( t \). Let \( Q \) be the set of all coalitional structures.

Suppose that \( i \in U^r_t \) and \( j \in U^c_t \) are matched according to (2.1). Abstracting away the details of the strategic interactions, we assume that the two players face a bargaining problem \( \langle V, v^0 \rangle \), where \( V \) is a compact convex subset of \( \mathbb{R}^2 \) and \( v^0 = (v^0_i, v^0_j) \) is the disagreement payoff vector.\(^3\)

Also abstracting away the details of the search process toward the agreeable outcome, we assume that a relation-specific pair of payoffs \( v = (v_i, v_j) \in V \) is drawn from \( V \) according to a probability measure \( \nu \) on \( V \), where \( v_i \) is the payoff for the row player \( i \) and \( v_j \) is the payoff for the column player \( j \). Since the probability distribution \( \nu \) restricts the class of feasible search technology, we impose only a mild regularity condition.

**Assumption 2.1.** Assume that for any non-empty open ball \( B \subset \mathbb{R}^2 \), \( B \cap V \neq \emptyset \), then \( \nu(B \cap V) > 0 \) holds.

This condition is imposed not to exclude any particular outcome ex ante sense. Other than this condition, \( \nu \) can be arbitrary so that our model can cover a broad class of search process between the two matched players.

Conditioned on \( v = (v_i, v_j) \), each agent chooses whether or not to agree to stay with the same partner in the next period. It is important to note that the decision of player \( i \) choosing \( R \) or \( A \) is conditioned only on \( v_i \) instead of \( (v_i, v_j) \).

The action space of each agent is \( \{A, R\} \) where “\( A \)” stands for “agree” and “\( R \)” for “not agree”. If both agents agree, then the two players remain matched for another round: \( \{i,j\} \in \mathcal{P}_{t+1} \).

If either player chooses \( R \), then both players return to the respective pools of singles, waiting for the next period for a new match: \( i \in U^r_{t+1} \) and \( j \in U^c_{t+1} \). If so, the bargaining fails and each player receives his own disagreement payoff. To simplify analysis, we assume that the disagreement payoff is conditioned only on the role rather than the identity of the player.

**Assumption 2.2.** \( \forall i \in \{1, \ldots, n\}, \forall j \in \{n + 1, \ldots, 2n\} \), \( v^0_i = \cdots = v^0_n = v^0_i \) and \( v^0_{n+1} = \cdots = v^0_{2n} = v^0_c \) for some \( (v^0_i, v^0_c) \in \mathbb{R}^2 \).

\(^3\)Our model is a generalization of [13], where only the workers are search for an outside option.
To simplify notation, we assume for the rest of the paper that 

\[ v^0_r = v^0_c = v^0. \]

The symmetry of the disagreement payoffs can be dropped without changing the main result of the paper.

If either player chooses \( R \) at time \( t \), then both return to the respective pools of singes, waiting for the next period for a new match, i.e., \( i \in U_{t+1}^r \) and \( j \in U_{t+1}^c \).

If both agents agree at time \( t \), then the two players remain matched for another period: 
\[ \{i, j\} \in \mathcal{P}_{t+1}. \]

At time \( t + 1 \), they obtain the same relation-specific payoffs \( v = (v_i, v_j) \) with probability \( \delta < 1 \). But, with probability \( 1 - \delta \), their relation-specific payoffs become anew and are drawn again from \( V \) according to \( \nu \).\(^4\) In either case, each player is given an opportunity to choose a response \( A \) or \( R \).

The timing of matches and decisions is illustrated in Figure 1.

---

**Figure 1. Timing of Matches and Decisions**

By a long term relationship, we mean a particular relationship that lasts for multiple periods.

**Definition 2.3.** We say agents \( i \) and \( j \) are in a long term relationship at time \( t \) if \( \exists k \geq 1 \) such that \( \forall t' \in \{t - k, \ldots, t\} \), 
\[ \{i, j\} \in \mathcal{P}_{t'}, \]

We do **not** assume that when a pair of players agrees upon a particular relationship, they are **committed** to the relationship. On the contrary, the players are allowed to terminate the relationship at any time after the relationship is formed. A long term relationship

\(^4\)This probability does not have to be the same as the one that governs the original draw as long as it satisfies the underlying assumptions.
is not assumed as in the repeated games, but is derived from social interactions in our model.

2.2. History and strategy. In the basic model, we assume that each player observes only his payoff and his own action in $t$:

$$s_{i,t} = (u_{i,t}, r_{i,t})$$

where $u_{i,t}$ is the payoff and $r_{i,t} \in \{A, R\}$ of player $i$ in period $t$. Let $s_t = (s_{1,t}, \ldots, s_{2n,t})$ be the profile of action in period $t$, which we call the social outcome in period $t$.

At the beginning of period $t$, player $i$ knows

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

which we call the private history of player $i$ in $t$. Let $H_{i,t}$ be the set of all private histories of player $i$ in $t$, and $H_i = \cup_{t \geq 1} H_{i,t}$ be the set of all private histories of player $i$.

A strategy of player $i \in I$ is a function

$$f_i : H_i \times \mathbb{R}^2 \rightarrow \{A, R\}.$$ 

Given a private history $h_{i,t}$ and a proposed payoff pair $q_t$, agent $i$'s action induced by $f_i$ is $f_i(h_{i,t}, q_t) \in \{A, R\}$. Let $\mathcal{F}_i$ be the set of strategies of agent $i$.

A strategy profile $f = (f_i)_{i \in I}$ induces a distribution over outcome paths. In period $t$, a social outcome is given by

$$s_t = ((s_{i,t})_{i \in I}, q_t),$$

where $q_t$ is the coalitional structure in period $t$. Let $h_t = (s_1, \ldots, s_{t-1})$ be the social history at time $t$.

Given $f = (f_1, \ldots, f_{2n})$, the payoff function of player $i$ is given by

$$U_i(f) = E^f \left[(1 - \beta) \sum_{t=1}^{\infty} \beta^{t-1} u_{i,t}\right]$$

where $E^f$ is the expectation operator induced by $f$, and $\beta \in (0, 1)$ is a discount factor.

2.3. Solution concept. The basic solution concept is Nash equilibrium.

**Definition 2.4.** $f$ is a Nash equilibrium if $\forall i, \forall f'_i \in \mathcal{F}_i$,

$$U_i(f) \geq U_i(f_{-i}, f'_i).$$

Conditioned on $h_i$, let $\mu_i(h_i)$ be the probability distribution over the information set reached by $h_i$. Let $\mu_i = (\mu_i(h_i))_{h_i \in H_i}$ and $\mu = (\mu_1, \ldots, \mu_{2n})$ be the system of beliefs. Given social history $h_t$, let us consider an extensive form game following $h_t$, whose game tree consists of nodes following $h_t$, and the probability distribution over the initial nodes is $\mu_i(h_{i,t})$ which is the belief of player $i$ conditioned on his private history $h_{i,t}$. We refer to this extensive form game as the continuation game after $h_t$, or simply, the continuation game.

Given private history $h_i$, define the continuation game strategy of player $i$ as

$$f_i(h'_i|h_i) = f_i(h_i \circ h'_i)$$

where $h_i \circ h'_i$ is the concatenation of $h_i$ and $h'_i$. Given history $h$, define $f(\cdot|h) = (f_1(\cdot|h), \ldots, f_{2n}(\cdot|h))$ as the profile of continuation game strategies.
Definition 2.5. \((f, \mu)\) is a sequential equilibrium or simply an equilibrium if \(\forall h, f(\cdot| h)\) constitutes a Nash equilibrium in the continuation game, where the probability distribution over the initial nodes of the continuation game is determined by \(\mu\), and \(\mu\) is computed by Bayes’ rule whenever possible.

Because every row player is ex ante identical, and so is every column player, a symmetric equilibrium is a natural focal point equilibrium.

Definition 2.6. An equilibrium \(f = (f_1, \ldots, f_{2n})\) is symmetric, if \(\exists f_r, f_c\) such that

\[
f_r = f_1 = \ldots = f_n \quad \text{and} \quad f_c = f_{n+1} = \ldots = f_{2n}.
\]

2.4. Stationary equilibrium. As a first step to compute an equilibrium, let us examine an equilibrium in which each player’s decision is conditioned on a small set of states rather than the entire history. To simplify exposition, we focus on the decision problem of a row player \(i\). A natural state would be the event whether player \(i\) is in the pool of singles \((i \in U)\), or whether player \(i\) forms a partnership with player \(j\) over payoff vector \((v_i, v_j)\). Let

\[
\Sigma_i = \{\emptyset\} \cup \{(j, v_i, v_j)| j \neq i, (v_i, v_j) \in V\}
\]

be the set of states of player \(i\), where \(\emptyset\) means \(i \in U\), and \((j, v_i, v_j)\) means \(i, j \in P\) after the two players agree to form a partnership over \(v\).

Let \(W_i(j, v_i, v_j)\) be the value function of player \(i\) in the steady state, conditioned on event that he is matched with player \(j\) around payoff vector \((v_i, v_j)\). Similarly, let \(W^0_i\) be the value function of player \(i\) in the steady state, conditioned on event that \(i \in U\).

We say that an equilibrium is a stationary equilibrium if each player’s strategy is stationary. Given a stationary equilibrium, the transition probability over the set \(Q\) of the coalitions clearly has a stationary distribution.

The model has a stationary equilibrium in which every player chooses \(R\) following any history. For the rest of the paper, we shall focus on an equilibrium which has a positive probability of forming a partnership.

Let us consider a stationary equilibrium in which the decision of player \(i\) is characterized by a threshold.

Definition 2.7. A strategy \(f_i\) is a stationary strategy if \(f_i\) is conditioned only on \(\Sigma_i\). A stationary strategy is a threshold strategy if \(\exists a_i\) such that player \(i\) chooses \(A\) if \(v_i > a_i\) and chooses \(R\) if \(v_i < a_i\). If \(f\) is an equilibrium in which every player uses a stationary strategy, then we call \(f\) a stationary equilibrium. If \(f\) is a stationary equilibrium in which every player uses a threshold strategy, then we call \(f\) a threshold equilibrium.

2.5. Existence and efficiency. As a benchmark, let us consider a symmetric threshold equilibrium in which each player chooses an action conditioned on a state in \(\Sigma_i\). Let us write subscript “\(r\)” and “\(c\)” for the representative row and column players, respectively, in place of \(i\) and \(j\). Let \(W_{r}(c, v_r, v_c)\) be the value function of a row player, conditioned on event that he is matched with a column player around payoff vector \((v_r, v_c)\) drawn from \(V\) according to probability measure \(\nu\). Similarly, let \(W^0_{r}\) be the value function of player \(r\), conditioned on event that the row player is in the pool of singles \(r \in U\).

Let us compute \(\{W^0_r, W_r(c, v_r, v_c); W^0_c, W_c(r, v_r, v_c)\}\) satisfying the equilibrium condition. We focus on the value functions of the row player, as the computation of the column
player’s value function follows the same logic. Conditioned on the event that \( r \in U^r \), it is optimal that player \( r \) agree \((v_r, v_c)\) if
\[
W_r(c, v_r, v_c) \geq W_r^0.
\]
Given the threshold decision rule, we can decompose \( W_r(c, v_r, v_c) \) and \( W_r^0 \) as
\[
W_r(c, v_r, v_c) = (1 - \beta)v_r + \beta \left[ \delta W_r(c, v) + (1 - \delta)W_r^0 \right],
\]
and
\[
W_r^0 = (1 - \beta)v_0 + \beta \int_{(v'_r, v'_c) \geq (W_r^0, W_c^0)} W_r(c, v')d\nu(v').
\]
Define
\[
p^{W_0} = \int_{(v'_r, v'_c) \geq (W_r^0, W_c^0)} d\nu(v')
\]
as the probability of the event that player \( i \) forms a partnership with a column player around some \( v' = (v'_r, v'_c) \). Let \( P_r \) be such an event. Solving (E.12) and (E.13), we have
\[
W_r^0 = \frac{(1 - \beta)\delta v_0 + \beta p^{W_0}E[v_r|P_r]}{1 - \beta \delta + \beta p^{W_0}}.
\]
Note that \( p^{W_0} \) and \( E[v_r|P_r] \) depend upon the entire profile of value functions. Thus, the existence proof of a stationary equilibrium requires to solve a fixed point problem. Let \( \overline{v} \) be the largest payoff any player can receive in \( V \). Clearly, \( v_0 < \overline{v} \).

**Lemma 2.8.** Given \( W_c^0 \), there exists a unique \( W_r^0 \) satisfying (2.6). A symmetric threshold equilibrium exists.

**Proof.** See Appendix A. \( \square \)

**Proposition 2.9.** A symmetric threshold equilibrium exists.

**Proof.** See Appendix B. \( \square \)

Let \((W_r^0, W_c^0)\) be the pair of value function in a symmetric threshold equilibrium conditioned on the event that a player is in the pool of singles. Given \((W_r^0, W_c^0)\), we can construct a stationary equilibrium strategy: If \( r \in U^r \), then player \( r \) chooses \( A \) only if
\[
v_r \geq W_r^0.
\]
From (2.6), we know this threshold strategy is indeed an optimal strategy, since
\[
v_r \geq W_r^0
\]
if and only if
\[
W_r(c, v_r, v_c) \geq W_r^0.
\]
It is straightforward to prove that the above action is optimal following every history. Furthermore, the same argument shows that after player \( i \) accepts \((v_r, v_c)\), then it is optimal for player \( i \) to choose \( A \). Thus, the existing partnership is dissolved in a threshold equilibrium only through the exogenous shock which arrives with probability \( 1 - \delta \) in each period.
By its nature, a stationary equilibrium does not entail a “punishment” phase, because each player does not discriminate his partner based upon what his partner did in previous rounds. On the other hand, because a social outcome may not be sustained by a myopic utility maximization of an individual player, it is necessary to encode a “punishment” into an equilibrium strategy. In a stationary equilibrium, each player has to choose his threshold to maximize his long run discounted average payoff, but also to deter his partner from deviating from the proposed equilibrium outcome. A remarkable feature of a stationary equilibrium is that as the players become more patient ($\beta \to 1$) and the continuation probability $\delta$ is close to 1, every player chooses a threshold in such a way that the social outcome becomes efficient.

**Proposition 2.10.** Let $P_V$ be the Pareto frontier of $V$ and $N_\epsilon(P_V)$ be the $\epsilon$ neighborhood of $P_V$. $\forall \epsilon > 0$, $\exists \delta', \beta'$ such that if $\delta \in (\delta', 1)$ and $\beta \in (\beta', 1)$, then $(W_r^0, W_c^0) \in N_\epsilon(P_V)$.

*Proof.* See Appendix C. $\Box$

We intentionally impose little restriction on $\nu$ in order to cover a broad class of search process. Without a specific restriction on $\nu$, it is difficult to narrow down where a stationary equilibrium outcome would be at the Pareto frontier. In fact, we can support any payoff vector in the Pareto frontier of $V$ as a limit point of a sequence of the profile of equilibrium payoffs by manipulating the probability distribution $\nu$ over $V$.

**Proposition 2.11.** Given $V$, take any Pareto efficient outcome $(x^*, y^*)$. Then for all $\zeta > 0$, there exist a measure $\nu$ on $V$ that satisfies the assumption in the main model and $\delta < 1$ such that for all $\delta \in (\delta, 1)$, there exists $(\bar{x}, \bar{y})$ such that $|\bar{x} - x^*| < \zeta$, $|\bar{y} - y^*| < \zeta$, $\Psi(\bar{x}, \bar{y}) = 0$, and $(\bar{x}, \bar{y})$ is a sink of the dynamics induced by the satisficing behavior.$^5$

*Proof.* See Appendix D. $\Box$

### 3. Equilibrium selection

While Proposition 2.10 delineates a possible long run social outcome, the same result needs to be strengthened in a number of ways. First, we justify that the restriction to the threshold rule is mild, especially if the number of players is large. That is, in any steady state outcome path of an equilibrium, the decision rule of each player is observationally equivalent to a slightly perturbed threshold rule. Second, as the probability distribution $\nu$ over $V$ models possible search process, we intentionally impose little restriction on $\nu$, in order to let our model cover a large class of search process toward an agreeable outcome. As a result, we have no information about

$$
\begin{bmatrix}
E[\nu_r|P_r] - W_r^0 \\
E[\nu_c|P_c] - W_c^0
\end{bmatrix}
$$

other than the fact that each component is non-negative. In order to narrow down the prediction of Proposition 2.10, we need more structure over $\nu$.

In subsection 3.1, we demonstrate that if the number of players is sufficiently large, the equilibrium play of each player can be approximated by a threshold rule. In subsection 3.3, we perturb the basic game by incorporating sympathy à la Hume, and prove that the

$^5$Note that this claim does not state that the above $(\bar{x}, \bar{y})$ is the unique absorbing state.
maxmin outcome is the only symmetric threshold equilibrium outcome of the perturbed game.

3.1. **Threshold rule, revisited.** Let take one step back from the class of stationary equilibria, and instead consider a larger class of equilibria that induce the outcome paths along which the behavior of the players are “steady.” We say that the outcome path is in a steady state, if the behavior of each player is stationary along the equilibrium path. Note that we impose little restriction on the equilibrium strategy off the equilibrium path.

As we regard the social outcome as a steady state of a social dynamics, it would be natural to look into the class of equilibria that entail a steady state. We say that a decision rule of player \(i\) is an \(\epsilon\) threshold rule if \(\exists a_i\) such that he agrees if \(v_i > W^0_i + \epsilon\) and disagrees if \(v_i < W^0_i - \epsilon\).

**Proposition 3.1.** Fix \(\beta, \delta\), an equilibrium \(f\) and its steady state. \(\forall \epsilon > 0, \exists N, \forall n \geq N,\) the decision rule of player \(i\) in the steady state is an \(\epsilon\) threshold rule with the threshold being \(W^0_i\).

**Proof.** See Appendix E. \(\square\)

3.2. **Estimating the value function.** Motivated by Proposition 3.1, let us imagine an agent whose strategy is characterized by a threshold rule, and uses the value function as the threshold. Since we are interested in the case where the players are patient and the probability of continuing the existing relationship is close to 1, the sample average

\[
a_{i,t} = \frac{1}{t} \sum_{k=0}^{t-1} u_{i,k}
\]

is a simple, yet sensible, estimator for the value function \(W^0_i\) in a steady state.

**Proposition 3.2.** Suppose that the economy is in a steady state. Then, for all \(\epsilon > 0\) there exist \(\beta < 1\) and \(\delta < 1\) such that for all \(\beta \in (\beta, 1)\) and all \(\delta \in (\delta, 1)\),

\[
\limsup_{t \to \infty} \left| W^0_i - \frac{1}{t} \sum_{k=0}^{t-1} u_{i,k} \right| \leq \epsilon
\]

with probability 1.

**Proof.** See Appendix F. \(\square\)

In fact, player \(i\)’s value function can be consistently estimated, if one observes his private history \(h_{i,t}\). Note that

\[
W^0_i = \frac{(1 - \beta \delta)v_0 + \beta p^{W^0} \mathbb{E}(v_i | (v_i, v_j) \geq (W^0_i, W^0_j))}{1 - \beta \delta + \beta p^{W^0}}.
\]

One can estimate \(p^{W^0}\) and \(\mathbb{E}(v_i | (v_i, v_j) \geq (W^0_i, W^0_j))\) from \(h_{i,t}\). \(p^{W^0}\) is the frequency that player \(i\) is matched with someone else. Thus, its sample counterpart would be

\[
\hat{p}_{i,t} = \frac{1}{t} \# \{ k \leq t | \exists j_k, \{ i, j_k \} \in \mathcal{P} \}. 
\]
Similarly, the sample counterpart of
\[ \int_{(v_i,v_j) \geq (W_i^0,W_j^0)} v_i d\nu \]
is
\[ \hat{v}_{i,t} = \frac{1}{t} \sum_{k=1}^{t} v_i \mathbb{1}(\exists j_k, \{i,j_k\} \in \mathcal{P}) \]
where \( \mathbb{1} \) is the indicator function. Then, the estimator for \( W_i^0 \) is
\[ (3.8) \quad \hat{W}_{i,t} = \frac{(1 - \beta \delta) v_0 + \beta \hat{v}_{i,t}}{1 - \beta \delta + \beta \hat{p}_{i,t}}. \]
If the equilibrium is an steady state, then the law of large numbers implies that
\[ \lim_{t \to \infty} \hat{W}_{i,t} = W_i^0. \]
It is important to note that the convergence holds under the assumption that the equilibrium path is in a steady state.

3.3. Perturbed game.

3.3.1. Information leakage. Although no player observes \( s_t \) at the end of period \( t \), each player knows his own action and the consequence of his action (which is summarized in terms of the payoff in the period), but also with a small probability, each player has an opportunity to observe the private history of the partner. That is, when two players are drawn from each pool of singles, each player’s strategy is “leaked” to the other party with a positive probability ([12]). This feature is to capture the idea that when two parties form a long term relationship, each party should have an opportunity to learn about the intention of the other party, if not perfectly, especially about the criterion to agree on a long term relationship.

To formulate the information leakage and the strategy, we need to specify the private history of each individual player. Each agent observes the outcome of his own match unless otherwise mentioned. Let \( \ell_{i,t} \) be the indicator function of an event that player \( i \) reveals
\[(s_{i,1}, \ldots, s_{i,t-1})\]
in period \( t \). If \( \ell_{i,t} = 1 \), then \( h_{i,t} \) is truthfully revealed to player \( j \) in period \( t \), who is matched to player \( i \). If \( \ell_{i,t} = 0 \), then no information about \( h_{i,t} \) is revealed to any player.

We assume that
\[ p_{\ell} = \mathbb{P}(\ell_{i,t} = 1) \quad \forall i, t \]
is a small positive number. Let \( \ell_{ij,t} = (\ell_{i,t}, \ell_{j,t}) \). Thus, the private history \( h_{i,t} \) of player \( i \) in period \( t \) is
\[(s_{i,1}, \ell_{ij_1,1}, \ldots, s_{i,t-1}, \ell_{ij_{t-1},t-1})\]
where \( j_k \) is the player who is matched to player \( i \) in period \( k \). Let \( H_i \) be the set of all private histories of player \( i \) in the games with information leakage. A strategy of agent \( i \in I \) is a function
\[ f_i : H_i \times \mathbb{R}^2 \to \{A, R\}, \]
measurable with respect to \( i \)’s information.
Given a private history $h_{i,t}$ and a proposed payoff pair $v_t$, agent $i$'s action induced by $f_i$ is $f_i(h_{i,t}, v_t) \in \{A, R\}$. Let $\mathcal{F}_i$ be the set of strategies of agent $i$.

3.3.2. Sympathy. The decision to initiate a long term relationship is based upon the selfish motive. However, we perturb the decision rule of the individual agent slightly so that each agent, albeit still selfish, evaluates the long term relationship from the viewpoint of the other agent with a small probability. Suppose that row agent $i$ and column agent $j$ are matched and draw $(v^r, v^c)$ as the payoff pair they receive. Suppose that with probability $p_\ell$, player $j$'s private history is leaked to player $i$.

Conditioned on the event, the row player can construct the estimator of player $j$'s value function and the threshold, which summarizes player $j$’s perception of the game.

**Definition 3.3.** Fix a small probability $\rho > 0$. Player $i$’s strategy $f_i$ is perturbed by sympathy, if with probability $1 - \rho p_\ell$, player $i$ chooses $A$ if $v_i > W^0_i$, but with probability $\rho p_\ell$, he chooses $A$ if $v_i > a_{j,t}$, where $a_{j,t}$ is a consistent estimator of $W^0_j$ in a steady state. We say that a strategy profile $f$ is perturbed by sympathy, if $vi, f_i$ is perturbed by sympathy. By a perturbed game, we mean the game obtained by perturb every player’s strategy by sympathy.

This perturbation captures an important aspect of sympathy in that one evaluate one’s future by putting oneself in the position of one’s partner. Yet, it is important to note that this type of sympathy does not inject the other player’s payoff into one’s payoff function, as in social preference assumed in behavioral economics ([1]). Rather, a player modifies his own threshold according to the other player’s threshold in order to evaluate his own payoff from the long term relationship. Through sympathy, we perturb the decision rule of a decision maker, but even after the perturbation, his own payoff remains the sole component of his objective function. Note that once one puts oneself in the position of one’s partner, the past average payoff of the partner becomes a consistent estimator of the future value.

3.4. Symmetric threshold equilibrium. We say $(r,c) \in V$ is strongly Pareto efficient, or efficient hereafter, if there is no $(r',c')$ such that $r' \geq r$ and $c' \geq c$ while at least one inequality holds with a strict inequality. We say that $(r^0,c^0)$ is a Rawlsian outcome if

$$\min(r^0,c^0) = \max \min_{(r,c) \in V} \{r,c\} \equiv R.$$  

If there are multiple Rawlsian outcomes, then some of them are not efficient. Define $(r^*,c^*)$ as the efficient Rawlsian outcome. We say $(r,c)$ is more equal than $(r',c')$ if

$$|r - c| \leq |r' - c'|.$$  

In particular, if $r = c$, then it is called the egalitarian outcome. Note that if $V$ is convex, then $(r^*,c^*)$ is the efficient Rawlsian outcome if and only if $(r^*,c^*)$ is the most egalitarian outcome among all efficient outcomes in $V$.

**Proposition 3.4.** Let us consider a perturbed game. $\forall \epsilon > 0$, $\exists \beta' \in (0,1)$, $\exists \delta' \in (0,1)$, $\forall \beta \in (\beta',1)$, $\forall \delta \in (\delta',1)$, if $(v^*_r,v^*_c)$ is the equilibrium expected payoff of a symmetric threshold equilibrium of a perturbed game, then $| (v^*_r,v^*_c) - (r^*,c^*) | < \epsilon$.  

**Proof.** See Appendix G. \qed
Proposition 3.4 is built on symmetric stationary equilibrium. While symmetric equilibrium strategy is often justified by the ex ante symmetry among players within the same population, it is not obvious why different players with different experience of the game has to follow the same strategy in the long run. We focus on stationarity for its simplicity, leaving unanswered the question about how a player ends up choosing his action based on the state rather than the details of his private history. Most importantly, Proposition 3.4 is built upon the presumption that each player can accurately compute the value function, which is affected by the actions of other players. In a society in which the private history is not revealed to another player, it is indeed a tall order for a player to infer the other player’s action in the continuation game. This section rigorously investigates the important questions left out by equilibrium analysis.

In the ensuing analysis, we write \( a_{i,t} \) as the threshold. We assume that \( a_{i,t} \) is an estimator of \( W_0^i \), without specifying whether \( a_{i,t} \) is a simple average of the past payoffs, or a more elaborate consistent estimator like (3.8). As we focus on the case where \( \beta, \delta < 1 \) are close to 1, the simple average is as good an estimator as a more elaborate estimator \( \hat{W}_0^i \).

Given the estimated value of \( W_0^i \), each agent tries to maximize his own expected discounted sum of the future payoffs, following the same reasoning process as in the equilibrium. Player \( i \) chooses to initiate a long term relationship if

\[
(4.9) \quad v_i > a_{i,t}.
\]

Note that (4.9) is derived from the equilibrium threshold rule. Thus, our players are very much forward looking, although they do not have the ability of perfectly foreseeing the future evolution of the outcome path. In this sense, our agents are boundedly rational but procedurally rational ([15]).

The long run behavior of player \( i \) is completely determined by the asymptotic properties of \( a_{i,t} \). Since we do not assume symmetry, we have to keep track of \( 2n \) dimensional vector \((a_{1,t}, \ldots, a_{2n,t})\) instead of a pair of thresholds. Given an initial condition \( a_0 \), let \( \Pi_t \) be the probability distribution of \( a_t \). A profile of average payoffs is said to be feasible if it is in the support of \( \Pi_t \). We write

\[
a^* = (r^*, \ldots, r^*, c^*, \ldots, c^*)
\]

where \((r^*, c^*)\) is the efficient Rawlsian outcome in \( V \).

Because the decision of agent \( i \) at \( t \) is determined by the threshold \( a_{i,t} \), the asymptotic property of \( \{(a_{1,t}, \ldots, a_{2n,t})\}_{t=1}^{\infty} \) is at the focus of our analysis. If \( V \) is symmetric, \( a^* \) coincides with the Nash bargaining solution, the egalitarian outcome and the maxmin outcome.

**Theorem 4.1.** \( \forall \varepsilon > 0, \forall \rho \in (0, 1), \exists \delta(\varepsilon) < 1, \forall \delta \in (\delta(\varepsilon), 1), \exists T_\delta \) such that \( \forall t \geq T_\delta \),

\[
\mathbb{P}(|a_t - a^*| \leq \varepsilon) \geq 1 - \varepsilon.
\]
Let
\[
    c_{i,T} = \frac{1}{T} \sum_{t=1}^{T} I(\exists j, \{i, j\} \in P)
\]
be the average frequency that agent \(i\) is in the long term relationship with someone else up to period \(T\). Then, \(\forall T \geq T_\delta\),
\[
P(c_{i,T} \geq 1 - \varepsilon \ \forall i) \geq 1 - \varepsilon.
\]

**Proof.** See Appendix H. \(\Box\)

The second part proves that the long term relationship becomes a pervasive social institution, as the frequency that each agent is in the long term relationship converges to 1. In order to prove the second part, we have to show that the behavior of different agents converges to the same behavioral rule which conforms to the egalitarian outcome, through a decentralized matching process.

To understand the precise role of the matching process and each condition in proving Theorem 4.1, let us consider the case of \(n = 1\) so that there are one row agent and one column agent who are repeatedly matched and then, separated with probability \(1 - \delta > 0\).

Choose \(\delta < 1\) sufficiently close to 1. The key difference from the “standard” repeated game is that each time when the two agents are “re-matched” they draw a new pair \((v_r, v_c) \in V\) and initiate the long term relationship only if \(v_i > a_i\) for \(i \in \{r, c\}\). If \(\delta < 1\) is close to 1, the two agents are engaged in the long term relationship most of the time, and the average payoff of each agent increases. Thus, \((a_r, a_c)\) converges to the Pareto frontier.

Suppose that \((a_r, a_c)\) is at the Pareto frontier, and \(a_r < a_c\). Because \((a_r, a_c)\) is at the Pareto frontier, \(\forall (v_r, v_c) \in V\), at least one party will find the outcome unsatisfying: either \(v_r < a_r\) or \(v_c < a_c\).

Each agent perturb his threshold by adopting his opponent threshold with a small probability. Thus, we need to consider 4 combinations of “actual” thresholds: \((a_r, a_c)\) with probability \((1 - \rho)^2\), \((a_r, a_r)\) with probability \(\rho(1 - \rho)\), \((a_c, a_c)\) with probability \(\rho(1 - \rho)\) and \((a_c, a_r)\) with probability \(\rho^2\).

Since \(a_c > a_r\), \((a_c, a_c)\) does not change the outcome: if \((v_r, v_c)\) does not induce a long term relationship under the original configuration \((a_r, a_c)\), no long term relationship is formed under \((a_c, a_c)\).

There is a positive mass of outcomes in \(V\) that used to fail to generate a long term relationship under \((a_r, a_c)\) but successfully forms a long term relationship under \((a_r, a_r)\). For those outcomes, the row agent’s payoff increases, who has the lower average payoff between the two.

Applying the same reasoning for any point at the Pareto frontier, we conclude that whenever
\[
    q = \min\{a_r, a_c\} < \min\{r^*, c^*\} = R^o,
\]
\[
    \hat{q} > 0.
\]
Whenever there is an agent whose average payoff from the past plays is below the Rawlsian social welfare, he can improve his long run payoff through the matching process. In a
certain sense, our matching process continuously improves the payoff the worst off agent in the society.

An important observation is that this aspect of the dynamics of $a_t$ can be generalized to a society with any number of agents. For a large finite $n \geq 1$, we shall focus on the evolution of the average payoff of the worst off agent in the society, regardless of his identity to show that his payoff converges to the Rawlsian social welfare.

5. Concluding Remarks

This paper develops a micro foundation for a social preference by approximating the social outcome as a steady state of a social dynamics. As we specify the details of the social dynamics and the perturbation of the underlying game, we can identify the dynamic foundation of a social preference. The same method opens up a new way to assess how plausible a social preference is by comparing the social dynamics and the kind of perturbations we need in order to sustain the social preference. In this paper, we focus on the maxmin criterion, which selects an egalitarian outcome whenever it is efficient, independently of the symmetry of the underlying social interactions. We admit a very broad class of search process by imposing little restriction on the probability distribution $\nu$ over $V$, while perturbing the strategy, motivated by a sympathetic behavior à la Hume. It would be an interesting exercise to see whether a social preference can be sustained by specifying a class of admissible search process, which can be modeled by restricting the admissible $\nu$ over $V$, without any perturbation over the strategy space.
Appendix A. Proof of Lemma 2.8

Fix $W_c^0$, and define

$$\varphi_r(w_r) = \frac{(1 - \beta \delta)v_0 + \beta p^W_0 \mathbb{E}[v_r | P_r]}{1 - \beta \delta + \beta p^W_0} - w_r.$$  

The first term is a convex combination of $v_0$ and $\mathbb{E}[v_r | P_r]$. In particular, if $w_r = v_0$, then $\mathbb{E}[v_r | P_r] \geq v_0$. Thus,

$$\varphi_r(v_0) \geq 0.$$  

On the other hand, if $w_r = \mathbb{E}$, then $p^W_0 = 0$, since there is no $v \in V$ such that $v_r \geq \mathbb{E}$ and $v_r \geq W_c^0$ simultaneously. Thus,

$$\varphi_r(\mathbb{E}) < 0.$$  

Since $\varphi_r(w_r)$ is a continuous function of $w_r$, $\exists W_c^0$ satisfying (2.6). A straightforward calculation shows that $\varphi(w_r)$ is a strictly increasing function at $w_r = W_c^0$ as long as $p^W_0 > 0$. Hence, if $p^W_0 > 0$, there can be at most one solution for (2.6). If $p^W_0 = 0$, then $W_c^0 = v_0$ is the only solution for (2.6).

Appendix B. Proof of Proposition 2.9

Consider a function over $V$, which maps $(w_r, w_c)$ to the unique solution $(w'_r, w'_c)$ which solves

$$\langle \varphi_r(w'_r), \varphi_c(w'_c) \rangle = (0, 0).$$  

Since this function is continuous over $V$ which is convex and compact, we have a pair of $(w_r, w_c)$ satisfying

$$\langle \varphi_r(w_r), \varphi_c(w_c) \rangle = (0, 0).$$  

This fixed point is the pair of value functions with desired properties.

Appendix C. Proof of Proposition 2.10

From (2.6), we know that $(W_c^0, W_r^0)$ must satisfy

(C.11)  

$$[1 - \beta \delta](1 - p^W_0)(v_0 - W_r^0) + \beta (1 - \beta)p^W_0 (\mathbb{E}[v_r | P_r] - W_r^0) = 0,$$

As $\beta \delta \to 1$, the first term vanishes in each equation. In order to maintain the equality,

$$p^W_0 (\mathbb{E}[v_r | P_r] - W_r^0) \to 0$$

and

$$p^W_0 (\mathbb{E}[v_c | P_c] - W_c^0) \to 0$$

which imply that $(W_c^0, W_r^0)$ must converge to the Pareto frontier. Thus, conditioned on the event that $(i, j) \in P$, the agreed payoff vector $(v_i, v_j)$ must be close to the Pareto frontier.

Appendix D. Proof of Proposition 2.11

The key idea is to construct $\nu$ in such a way that the selected outcome is a limit point of the social dynamics. Choose a Pareto efficient outcome $(x^*, y^*)$ in $V$. Fix $\zeta > 0$. Then construct a measure $\nu$ on $V$ in such a way that there is a huge mass in, say, $\frac{1}{\zeta} \zeta$-neighborhood of $(x^*, y^*)$ with a positive support everywhere on $V$. We can ensure that $\nu$ has a continuous density function uniformly bounded away from 0 over its support.

Suppose that every row agent has the average payoff of $x$, while every column agent has the average payoff of $y$. Take any agent $i \in I'$. The corresponding direction in which the average payoff of this agent moves is given by

$$\frac{1 - \delta}{1 - \delta + p_i} \left[ v_i^* - x \right] + \frac{p_i}{1 - \delta + p_i} \left[ \mathbb{E}[u_i | u_i > x, u_j > y, j \in I'] - x \right].$$  

Similarly, if we take any agent $j \in I'$, the direction is given by

$$\frac{1 - \delta}{1 - \delta + p_j} \left[ v_j^* - y \right] + \frac{p_j}{1 - \delta + p_j} \left[ \mathbb{E}[u_j | u_i > x, u_j > y, j \in I'] - y \right].$$
Since all of the row (column) agents have a common average payoff, \( p_i = p_j \) holds for all \( i \in I \). Letting \( \delta \) go to one, we can find a point \((\bar{x}, \bar{y})\) or “near” the line segment connecting \((v'_i, v'_j)\) and \((x^*, y^*)\) where the first term and the second term are balanced. Moreover, it is verified that this point \((\bar{x}, \bar{y})\) is absorbing by the construction of \( \nu \).

**Figure 2.** Need more restriction on \( \nu \)

**Appendix E. Proof of Proposition 3.1**

Without loss of generality, let us assume that player \( i \) is a row player, whose value function in the steady state must satisfy

\[
W_i(j,v) = (1 - \beta)v_i + \beta \left[ \delta W_i(j,v) + (1 - \delta)W_i^0 \right]
\]

(\text{E.12})

\[
W_i^0 = (1 - \beta)v_0 + \beta \left[ (1 - p^{W_0})W_i^0 + \int_{(v'_i, v'_j) \geq (W_i^0, W_j^0)} W_i(j,v) \, dv(v) \right]
\]

(\text{E.13})

where

\[
p^{W_0} = \int_{(v'_i, v'_j) \geq (W_i^0, W_j^0)} dv(v) = P((v'_i, v'_j) \geq (W_i^0, W_j^0)).
\]

Suppose that \( \exists h, \exists j, \exists v = (v_i, v_j) \) such that

\[
v_i > W_i^0 + \epsilon
\]

but \( f_i(h, (v_i, v_j)) = R \). Since \( W_i^0 \geq v_0, v_i > v_0 \). Since \( R \) is an equilibrium response,

\[
(1 - \beta)v_i + \beta \left[ \delta W_i(j,v) + (1 - \delta)W_i^0 \right] \leq (1 - \beta)v_0 + \beta W_i^0
\]

(\text{E.14})

must hold, where \( \bar{W}_i(j,v) \) and \( \bar{W}_i^0 \) are the respective values of agent \( i \) when he deviates to choose \( A \) instead of \( R \).
We compute the bound for \( \bar{W}_i(j,v) - W_i(j,v) \) and \( \bar{W}_i^0 - W_i^0 \). Let

\[
\Delta v_{\text{max}} = \max_i \max_{(v_i,v_j) \in \mathcal{V}} |v_i - v_j|.
\]

be the maximum difference of one period payoff of any player \( i \).

Being in a steady state, player \( j \neq i \) would not take an action different from what he would have done, if player \( i \) followed the equilibrium action \( R \). The maximum payoff loss is bounded by \( \Delta v_{\text{max}} \). However, such an event can arise only if player \( j \) is met with probability \( 1/n \), and if player \( i \)'s private history is leaked with probability \( p_t \). Thus, for a given \( \beta, \delta, p_t \), if \( n \) is sufficiently large, the deviating player \( i \) expects others to play their steady state actions with a large probability for a long time.

Combining these arguments, we have

\[
\max \left( |\bar{W}_i(j,v) - W_i(j,v)|, |\bar{W}_i^0(j,v) - W_i^0| \right) \leq \frac{\Delta v_{\text{max}}}{1 - \delta} \sum_{t=1}^{[\log_2 n]} \beta^t\frac{\max_{j,v} |v_j - v_j'|}{\beta}\]

if \( \beta \neq 1/2 \), where \( [x] \) is the largest integer among those smaller than \( x \). The right hand side converges to 0 as \( n \) goes to infinity. The case of \( \beta = 1/2 \) is dealt with in a similar manner. Since \( v_i > v_0 \), for any \( \epsilon > 0 \), there exists \( N \) such that \( n > N \) implies

\[
(E.15) \quad v_i \leq W_i^0 + \epsilon,
\]

which is a contradiction.

**Appendix F. Proof of Proposition 3.2**

As a benchmark, suppose that in a steady state, (E.12) and (E.13) must hold, which implies that \( W_i^0 \) is equal to

\[
(F.16) \quad \frac{1 - \beta \delta}{1 - \beta \delta + \beta p_i W_0} v_i^0 + \frac{\beta}{1 - \beta \delta + \beta p_i W_0} \int_{(v_i,v_j) \geq (W_i^0,W_j^0)} v_i d\nu.
\]

Because of a finite number of players, the value function is affected by the action by an individual player. Yet, for a large \( n \), this impact diminishes. Following the proof of Proposition 3.1, we have \( \forall \epsilon > 0, \exists N \) such that \( \forall n \geq N \),

\[
(F.17) \quad \left| W_i^0 - \frac{1 - \beta \delta}{1 - \beta \delta + \beta p_i W_0} v_i^0 + \frac{\beta}{1 - \beta \delta + \beta p_i W_0} \int_{(v_i,v_j) \geq (W_i^0,W_j^0)} v_i d\nu \right| < \frac{\epsilon}{3}.
\]

By Proposition 3.1, the transition probability from the pool of singles to a matched pair is “approximately”

\[
p_i W_0
\]

while the transition from a matched pair to a pool of singles is \( 1 - \delta \). Thus, in a steady state, the long run expected payoff of player \( i \) is approximated by

\[
(F.18) \quad \frac{1 - \delta}{1 - \delta + p_i W_0} v_i^0 + \frac{1}{1 - \delta + p_i W_0} \int_{v \geq W_0} v_i d\nu,
\]

which is the expected payoff if player \( i \) uses a threshold rule and the equilibrium threshold is \( W_i^0 \) in a steady state.

For a fixed \( \epsilon > 0 \), we first choose \( \beta < 1 \) and \( \delta < 1 \) close to 1 so that (F.16) and (F.18) is within \( \epsilon/3 \). Then, choose \( n \) sufficiently large that (F.17) holds to have

\[
|W_i^0 - EU_{i,t}| < \frac{2\epsilon}{3}.
\]
We replace \( E_{u_{i,t}} \) by its sample counterpart
\[
\frac{1}{l} \sum_{k=0}^{l-1} u_{i,k}
\]
and invoke the law of large numbers to prove the conclusion of the proposition.

**Appendix G. Proof of Proposition 3.4**

Let us write \( P_{x,y} \) where \( x, y \in \{r, c\} \) as the state in which a row player uses \( W^0_r \) and a column player uses \( W^0_c \). If \( x \neq r \), then we say that the row player is sympathetic to the column player and similarly, if \( y \neq c \), then the column player’s strategy is perturbed by sympathy which occurs with probability \( \rho > 0 \).

Define an event
\[
P_\gamma = \cup_{x,y \in \{r,c\}} P_{x,y}
\]
and
\[
\gamma = P(P_\gamma).
\]
Recall that \( \gamma > 0 \) is fixed and \( \nu \) which has a full support over \( V \). If
\[
\min(W^0_r, W^0_c) < \min(r^*, c^*)
\]
where \((r^*, c^*)\) is the efficient maxmin outcome, then
\[
P(P_\gamma) > 0.
\]
In particular, if \( W^0 = \min(W^0_r, W^0_c) < \min(r^*, c^*) \), then
\[
P(P_{rr}) > 0.
\]

To prove the proposition by way of contradiction, suppose that there is \( \epsilon > 0 \), a sequence \( \beta_n \delta_n \to 1 \) and the associated value functions \((W^0_{r,n}, W^0_{c,n})\) such that
\[
\liminf_{n \to \infty} |(W^0_{r,n}, W^0_{c,n}) - (r^*, c^*)| \geq \epsilon.
\]
Let \((W^0_r, W^0_c)\) be a limit point of the sequence. Without loss of generality, suppose that
\[
W^0_r = \min(W^0_r, W^0_c) < \min(r^*, c^*).
\]
A simple calculation shows that
\[
W^0_{r,n} = \frac{(1 - \beta_n \delta_n) v^0 + \beta_n \gamma_n E(V_r|P_\gamma)}{1 - \beta_n \delta_n + \beta_n \gamma_n}
\]
where
\[
\gamma_n = P(P_\gamma)
\]
associated with \((W^0_{r,n}, W^0_{c,n})\).

By assumption, \( \beta_n \delta_n \to 1 \). Thus, in order to have the equality hold,
\[
E(v_r|P_\gamma) - W^0_{r,n} \to 0
\]
or
\[
\gamma_n \to 0.
\]
However, since \((W^0_{r,n}, W^0_{c,n}) \to (W^0_r, W^0_c)\) which is away from the efficient Rawlsian outcome,
\[
\liminf_{n \to \infty} E(v_r|P_\gamma) - W^0_{r,n} > 0
\]
or
\[
\liminf_{n \to \infty} \gamma_n > 0.
\]
This contradiction proves the proposition.
APPENDIX H. PROOF OF THEOREM 4.1

Our focus of interest is to understand the asymptotic properties of $a_t$. We first use the stochastic approximation technique, analyzing the dynamics of the mean of the average payoffs of individual agents instead of the sample path induced by the stochastic process. Instead of analyzing the entire profile of $a_t$, we focus on a key component in $a_t$ to pin down the stability and the convergence properties of $a_t$.

H.1. Preliminaries. In order to make this paper self-contained, we introduce notation and define concepts needed for later analysis. We draw our material from [10], which provides us with a comprehensive introduction to the stochastic approximation technique.

We write the simple average of the past payoffs in a recursive form

\[(H.19) \quad a_{i,t} = a_{i,t-1} + \frac{1}{T}(u_{i,t-1} - a_{i,t-1}),\]

where $u_{i,t-1}$ is agent $i$'s utility realized in period $t-1$ ($i \in I$). Let $a_t = (a_{1,t}, \ldots, a_{2n,t})$. The following conditions are needed. In these conditions, $a$ means a fixed feasible average payoff profiles rather than the abbreviation of $a_t$.

A.1 $\sup_{i,t} E|u_{i,t}|^2 < \infty$

A.2 For each $i \in I$, there is a measurable function $\hat{\Psi}_i(a_t, (u_{t-1}, U_t, P_t))$ such that

\[E(u_{i,t}) = E[u_{i,t}|u_s, s < t; (U_s, P_s), s \leq t] = \hat{\Psi}_i(a_t, (u_{t-1}, U_t, P_t)).\]

A.3 Given a fixed $a \in \mathbb{R}^I$, there exists $\Psi_i(a)$ such that

\[\lim_{t \to \infty} \lim_{T \to \infty} \frac{1}{T} \sum_{\tau=t}^{t+T} E_t \left[\hat{\Psi}_i(a, (u_{\tau}, U_{\tau}, P_{\tau}))\right] = \Psi_i(a)\]

with probability 1 for each $i \in I$ and for any given feasible $a$.

A.4 $\Psi_i(a)$ is continuous in $a$.

Define $\tau_0 = 0$, and $\tau_K = \sum_{t=1}^{K} 1/t$ for $K \geq 1$. For $\tau > 0$, define $m(\tau)$ as the unique $K$ such that $\tau_K \leq \tau < \tau_{K+1}$. Given a discrete time stochastic process $\{a_t\}_{t=1}^{\infty}$, we construct a continuous time stochastic process $a^0(\tau)$ through a continuous time interpolation:

\[a^0(\tau) = a_K \quad \text{if} \quad \tau_K \leq \tau < \tau_{K+1}.\]

Define the (left) shifted process $a^K(\tau)$ as

\[a^K(\tau) = a^0(\tau_K + \tau).\]

The classic results of the stochastic approximation consist of two parts: the approximation and the convergence of $\{a_t\}$. Let us state a basic theorem (Theorem 2.1 in [10]) adapted for our model. A complete proof for much more general environments can be found in [10].

Theorem H.1. Consider an ordinary differential equation (ODE)

\[\dot{a}_t = \Psi_i(a) \quad \forall i\]

or more compactly,

\[(H.20) \quad \dot{a} = \Psi(a)\]

where $\Psi = (\Psi_1, \ldots, \Psi_{2n})$.

Suppose that A.1-A.4 hold for (H.19). Then, the sample path of the discrete time stochastic process $\{a_t\}_{t=1}^{\infty}$ is approximated by a trajectory of (H.20) for a large $t$:

\[(H.21) \quad \lim_{K \to \infty} \lim_{\tau \to \infty} \left[ a^K_i(\tau) - a^K_i(0) - \int_0^\tau \Psi_i(a(s))ds \right] = 0 \quad \forall i\]

with probability 1 where $a(0) = a^K(0)$. 
Proof. Hence, there is a unique stationary distribution from the previous play, and that of all the states.\footnote{Note that \( u_t \) depends on \( u_{t-1} \) as well as \((U_t, \mathcal{P}_t)\). However, since \((U_t, \mathcal{P}_t)\) depends only on \((U_{t-1}, \mathcal{P}_{t-1})\), we can focus on the transition of \((U_t, \mathcal{P}_t)\).}

Let (1) be an agent who has the lowest average payoff among \( \{a_1, \ldots, a_{2n}\} \):
\[
(a_1) = \min\{a_1, \ldots, a_{2n}\}.
\]

A state is, therefore, described as a pair \((U; \mathcal{P})\). Note that \( \mathcal{P} \) is empty if no pair is in the satisfied match from the previous play, and that \( U \) is empty if everyone is matched and satisfied. Let \( Q \) be the collection of all the states.

For a fixed \( a \), the matching technology induces a Markov process over \( Q \). Let \( \Gamma_a \) be the transition matrix. A probability distribution \( \mu^*(a) \) over \( Q \) (a row vector) is a stationary distribution if it solves the balance equation
\[
(H.22) \quad \mu^*(a) \Gamma_a = \mu^*(a).
\]

Proposition H.2. For all \( a \in \mathbb{R}^l \), there exists a unique stationary distribution \( \mu^*(a) \), which is a continuous function of \( a \). Consequently, A.3 and A.4 hold.

Proof. We first show the unique existence of \( \mu^*(a) \) \( \forall a \). To this end, let us first consider the case where \( B(\underline{a}, \underline{a}) \cap V \) has non-empty interior:
\[
\text{int}\{B(\underline{a}, \underline{a}) \cap V\} \neq \emptyset.
\]

In this case, since every agent has an opportunity to sample \( \underline{a} \), any row agent has a chance to form a long term relationship with any column agent with a positive probability. Moreover, every state, including \((I; \emptyset)\), can be reached from \((I; \emptyset)\) in one step, and vice versa. Thus, the Markov process on \( Q \) is ergodic. Hence, there is a unique stationary distribution \( \mu^*(a) \). Let
\[
\Psi_i(a) = E_{\mu^*(a)} u_i.
\]

Using a property of the Markov chain, we can conclude that
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E_i u_i = \Psi_i(a).
\]

Next, suppose that
\[
\text{int}\{B(\underline{a}, \underline{a}) \cap V\} = \emptyset
\]
holds. Then no \( i \in U^r \) or \( j \in U^c \) can form a long term relationship, while the existing long term relationship breaks down with probability \( 1 - \delta \) in each period. Thus, the only recurrent state is \((I; \emptyset)\), where \( \mu^*(a) \) is concentrated.

Let (1) be an agent who has the lowest average payoff among \( \{a_1, \ldots, a_{2n}\} \):
\[
(a_1) = \min\{a_1, \ldots, a_{2n}\}.
\]

A state is, therefore, described as a pair \((U; \mathcal{P})\). Note that \( \mathcal{P} \) is empty if no pair is in the satisfied match from the previous play, and that \( U \) is empty if everyone is matched and satisfied. Let \( Q \) be the collection of all the states.

For a fixed \( a \), the matching technology induces a Markov process over \( Q \). Let \( \Gamma_a \) be the transition matrix. A probability distribution \( \mu^*(a) \) over \( Q \) (a row vector) is a stationary distribution if it solves the balance equation
\[
(H.22) \quad \mu^*(a) \Gamma_a = \mu^*(a).
\]

Proposition H.2. For all \( a \in \mathbb{R}^l \), there exists a unique stationary distribution \( \mu^*(a) \), which is a continuous function of \( a \). Consequently, A.3 and A.4 hold.

Proof. We first show the unique existence of \( \mu^*(a) \) \( \forall a \). To this end, let us first consider the case where \( B(\underline{a}, \underline{a}) \cap V \) has non-empty interior:
\[
\text{int}\{B(\underline{a}, \underline{a}) \cap V\} \neq \emptyset.
\]

In this case, since every agent has an opportunity to sample \( \underline{a} \), any row agent has a chance to form a long term relationship with any column agent with a positive probability. Moreover, every state, including \((I; \emptyset)\), can be reached from \((I; \emptyset)\) in one step, and vice versa. Thus, the Markov process on \( Q \) is ergodic. Hence, there is a unique stationary distribution \( \mu^*(a) \). Let
\[
\Psi_i(a) = E_{\mu^*(a)} u_i.
\]

Using a property of the Markov chain, we can conclude that
\[
\lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E_i u_i = \Psi_i(a).
\]

Next, suppose that
\[
\text{int}\{B(\underline{a}, \underline{a}) \cap V\} = \emptyset
\]
holds. Then no \( i \in U^r \) or \( j \in U^c \) can form a long term relationship, while the existing long term relationship breaks down with probability \( 1 - \delta \) in each period. Thus, the only recurrent state is \((I; \emptyset)\), where \( \mu^*(a) \) is concentrated.
We conclude that for all \(a\), there exists a unique \(\mu^*(a)\) solving the balance equation (H.22). Note that \(\Gamma_a\) is a continuous function of \(\alpha\) since \(\nu\) has no mass point. Therefore, \(\mu^*(a)\) is continuous in \(\alpha\), which leads to A.3 of Subsection H.1. \(\square\)

Given a state \((U, \mathcal{P})\), and given \(i \in U^c\) and \(j \in U^c\), let \(p_{ij}(a)\) be the probability that agents \(i\) and \(j\) form a long term relation upon matching, which is calculated as

\[
p_{ij}(a) = \sum_{k \in l} \sum_{t \in I} p_k^t p_i^t \nu(B(a_k, a_t)) \geq m^2 \sum_{k \in l} \sum_{t \in I} \nu(B(a_k, a_t))
\]

where

\[
m = \min_{\nu, k' \in I} p_{k'}^t > 0.
\]

The probability \(p_{ij}(a)\) for \(i \in U^c\) and \(j \in U^c\) is defined in a similar manner. Since \(\nu(\cdot) \geq 0\) always holds, (H.23) implies

\[
p_{ij}(a) \geq m^2 \nu(B(a, a)), \quad \forall i, j \in I.
\]

This stationary distribution \(\mu^*(a)\) determines the unconditional probability \(z_i(a)\) that agent \(i \in I\) remains in the pool of singles:

\[
z_i(a) = \sum_{U \ni i} \mu^*(a)(U; \mathcal{P}), \quad \forall i \in I.
\]

We can define the probability of his forming a (new) long term relation with some column agent conditioned on \(i \in U^c\), which is denoted by \(p_i(a)\):

\[
p_i(a) = \frac{\sum_{j \in I^c} \sum_{U \ni j} \mu(a)(U, \mathcal{P})}{\#U/2} p_{ij}(a) \geq m^2 \nu(B(a, a)),
\]

where we make use of (H.24) for the inequality. We can also define the corresponding probability for a column agent in a similar manner.

Since the probability of returning to the pool is 1 - \(\delta\) after both agree to continue the relationship,

\[
z_i(a) = \frac{1 - \delta}{1 - \delta + p_i(a)} \leq \frac{1 - \delta}{1 - \delta + m^2 \nu(B(a, a))}.
\]

Recall that (1) is the agent whose average payoff is the lowest among \(\{a_1, \ldots, a_2n\}\). If (1) is a row agent, we have

\[
\hat{a} = \mathbb{E}_{\mu^*(a)}u(1) - a
\]

\[
= \left[ z(1)(a) v_0 + (1 - z(1)(a)) \sum_{j \in I^c} \sum_{k \in l} p_k^t p_i^t \int_{B(a_k, a_t)} u^r d\nu \right] - a
\]

\[
\geq \left[ z(1)(a) v_0 + (1 - z(1)(a)) - m^2 a + m^2 \int_{B(a_k, a_t)} u^r d\nu \right] - a
\]

\[
\geq z(1)(a)(v_0 - a) + m^2 \int_{B(a_k, a_t)} u^r d\nu - a
\]

where the inequality in (H.28) holds since \(\int_{B(a_k, a_t)} u^r d\nu \geq a \nu(B(a_k, a_t))\) holds for all \(k, \ell \in I\).

Then we have the following key result.

**Lemma H.3.** \(\forall \zeta > 0, \ \exists \varepsilon > 0, \ \exists \hat{a} < 1 \ \forall \delta \in (\hat{a}, 1)

\[
\hat{a} < R^\varepsilon - \zeta \Rightarrow \hat{a} \geq \varepsilon.
\]

where

\[
R^\varepsilon = \min\{r^*, c^*\}
\]

is the Rawlsian social welfare.
Proof. Fix $\zeta > 0$. Consider
\[ G(b) = \frac{\int_{(u^*,w^*) \in B(b,b)} u^* d\nu}{\nu(B(b,b))} - b \]
defined for $b \in \left[ \min_{(u^*,w^*) \in V} \min \{ u^*, w^* \}, r^* - \zeta \right]$.

Since $\nu$ has no atom, $G(b)$ is continuous, which represents the expected gain from forming a long term relationship. Since $b \leq r^* - \zeta$, the expected gain is strictly positive. Thus,
\[ D = \inf_{b \in \left[ \min_{(u^*,w^*) \in V} \min \{ u^*, w^* \}, r^* - \zeta \right]} G(b) > 0 \]
which depends only on $\zeta$. Thus, we have
\[ \hat{a} \geq z_{(1)}(a)(v_0 - a) + m^2 D. \]

Since $v_0 - a < 0$ may hold, we need to show $z_{(1)}(a)(v_0 - a)$ is sufficiently close to 0 in order to prove that the right hand side is positive.

Note that
\[ \nu(B(a,a)) \geq \nu(B(r^* - \zeta, r^* - \zeta)) > 0 \]
holds for any $a$ with $a < r^* - \zeta$. From (H.27), $z_{(1)}(a)$ can be made arbitrarily small by letting $\delta \to 1$. Therefore, $\exists \delta < 1$ such that $\forall \delta \in (\delta, 1)$,
\[ z_{(1)}(a)(v_0 - a) > -\frac{1}{2} m^2 D \]
for all $a$ with $a < r^* - \zeta$.

Choose $\varepsilon < \frac{1}{2} m^2 D$, and take any $\delta \in (\hat{a}, 1)$. Then, we have $\hat{a} > \varepsilon$ for any $a$ satisfying $a < r^* - \zeta$. \qed

H.3. Convergence. Let us state the main convergence theorem.

Theorem H.4. Suppose that $(r^*, c^*) \in V$ is the efficient Rawlsian outcome. \forall $\varepsilon > 0$, \exists $T$ such that $\forall t \geq T$
\[ P(|a_t - a^*| < \varepsilon) \geq 1 - \varepsilon. \]

The proof takes a number of steps. To this end, we represent $(r^*, c^*)$ into analytically more convenient form. First, note that $(r^*, c^*)$ must be located at the Pareto frontier of $V$, which is assumed to be convex and compact.

Since $(r^*, c^*)$ is in the Pareto frontier, there is an outer norm $(\gamma_1, \gamma_2)$ such that $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$ as depicted in Figure 3.

We can normalize the outer norm so that
\[ \gamma_1 + \gamma_2 = 1. \]

Since $V$ is convex, there may be multiple outer norms at $(r^*, c^*)$ if the boundary of $V$ is not smooth at $(r^*, c^*)$. To handle the non-differentiable case, we need to select a particular outer norm.

If $r^* < c^*$, then define
\[ \gamma^* = \sup \{ \gamma_1 \geq 0 \mid (\gamma_1, \gamma_2) \text{ is an outer norm at } (r^*, c^*) \}. \]

If $r^* > c^*$, then define
\[ \gamma^* = \inf \{ \gamma_1 \geq 0 \mid (\gamma_1, \gamma_2) \text{ is an outer norm at } (r^*, c^*) \}. \]

If $r^* = c^*$, then choose $\gamma^* = \gamma_1$ for any $(\gamma_1, \gamma_2)$ which is an outer norm at $(r^*, c^*)$. Since $V$ is convex, $(r^*, c^*)$ is the utilitarian outcome if the weight assigned to the row agent’s payoff is $\gamma^*$:
\[ (r^*, c^*) \in \arg \max_{(r,c) \in V} \gamma^* r + (1 - \gamma^*) c. \]

Applying [10], we approximate the sample path of $a_t$ by the trajectory of the following ODE: $\forall i \in I^r$
\[ \dot{a}_i = E_{a^*}(a_i)u_i - a_i \]
\[ = \left[ z_i(a)(v_{0i} + (1 - z_i(a)) \sum_{j \in I^r} \sum_{k \in I^r} \sum_{r \in I} p_k p_j \frac{\int_{(u^*,w^*) \in B(a_k,a_i)} u^* d\nu}{\nu(B(a_k,a_i))} \right] - a_i \]
\[ \text{H.31} \]
and $\forall j \in I^c$

(H.32) \[
\dot{a}_j = E_{\mu^*(a)} u_j - a_j = \left[ z_j(a) u_0^c + (1 - z_j(a)) \sum_{i \in I^r} \sum_{k \in I} \sum_{l \in I} p^j_{ikl} \frac{\int_{(u^r,a^c) \in B(a_k,a_{2n})} u^d \nu}{\nu(B(a_k,a_{2n}))} \right] - a_j
\]

where \[B(a_i,a_j) = \{(v_r,c_c) \in V \mid a_i \leq v_r, \ a_j \leq c_c\}.$

It suffices to show that this system of ODE converges to $a^*$. From now on, by $a_t$ or $a$, we mean the deterministic trajectory induced by ODE rather than the original stochastic process of the average payoffs.

Without loss of generality, we have assumed that $r^* \leq c^*$, as the other case where $r^* \geq c^*$ follows from the symmetric argument. Depending upon the marginal rate of substitution at $(r^*,c^*)$, we consider two cases: $0 < \gamma^* < 1$ and $\gamma^* = 1$. The remaining case $\gamma^* = 0$ follows from the case of $\gamma^* = 1$.

H.3.1. $0 < \gamma^* < 1$. In this case, $r^* = c^*$. By Lemma H.3, we know that for a sufficiently large $t$, $a$ must be close to $R^o = r^* = c^*$. Let us write its direct implication in a form more convenient for our proof.

Lemma H.5. $\forall \zeta > 0$, $\exists T > 0$ such that $\forall t \geq T$,

\[a \geq r^* - \zeta = c^* - \zeta.\]

Let \[\bar{a} = \max\{a_1, \ldots, a_{2n}\}\]

and $(2n)$ be the agent whose average payoff is $\bar{a}$, i.e.,

\[a_{(2n)} = \bar{a}.\]

\[
\begin{align*}
\text{Figure 3. Case of } 0 < \gamma^* < 1
\end{align*}
\]
Lemma H.6. If $0 < \gamma^* < 1$, then for all $\zeta > 0$, there exists $T$ such that for all $t \geq T$,

\[
\bar{a} \leq r^* + \zeta \left[ \max \left( \frac{1}{\gamma^*}, \frac{1}{1 - \gamma^*} \right) \right].
\]

Proof. Take $\zeta > 0$ as given. Suppose $a \geq r^* - \zeta$. Suppose also

\[
\bar{a} > r^* + \zeta \left[ \max \left( \frac{1}{\gamma^*}, \frac{1}{1 - \gamma^*} \right) \right]
\]

Suppose that $(2n) \in I^c$, i.e., $(2n)$ is a row agent. Then $(u^r, u^c) \in V$ and $u^r \geq a$ imply that

\[ u^r \leq r^* + \zeta \left[ \max \left( \frac{1}{\gamma^*}, \frac{1}{1 - \gamma^*} \right) - 1 \right]. \]

See Figure 3. Note also that $\bar{a} > r^* + \zeta > v^c_0 + \zeta$ holds. Hence, in the light of (H.31), we have

\[ \bar{a} = a_{(2n)} < -\zeta. \]

Thus, there exists $T$ such that for all $t \geq T$

\[ a_{(2n)} \leq r^* + \zeta \left[ \max \left( \frac{1}{\gamma^*}, \frac{1}{1 - \gamma^*} \right) \right]. \]

A similar argument holds for $(2n) \in I^c$. Thus, there exists $T$ such that for all $t \geq T$, (H.33) holds, as desired.

Combining Lemmata H.5 and H.6 with Theorem H.1, we obtain the desired convergence result.

Theorem H.7. Suppose that $0 < \gamma^* < 1$ holds. Then, for all $\zeta > 0$, there exists $\delta < 1$ such that for all $i \in I$, $r^* - \zeta = \liminf_{t \to \infty} a_{i,t} \leq \limsup_{t \to \infty} a_{i,t} \leq r^* + \zeta = c^* + \zeta$ holds with probability 1 for all $i \in I$.

H.3.2. $\gamma^* = 1$ or $\gamma^* = 0$. We only examine the case of $\gamma^* = 1$. The remaining case follows by the same logic. If $\gamma^* = 1$, then

\[ r^* = \max_{(r,c) \in V} r \leq c^*. \]

In this case, the row agents’ payoffs are bounded from above by $r^*$. Therefore, Theorem H.7 continues to hold for row agents, i.e., $\forall i \in I^c$, $\exists \delta < 1 \forall \delta \in (\delta, 1)$

\[ r^* - \zeta \leq \liminf_{t \to \infty} a_{i,t} \leq \limsup_{t \to \infty} a_{i,t} \leq r^* + \zeta, \quad \forall i \in I^c = \{1, \ldots, n\}, \]

with probability 1.

We use the strict convexity of $V$ to characterize the bound for the column agents’ average payoffs.

Proposition H.8. Suppose that $\gamma^* = 1$ holds. Then, $\forall \xi > 0 \exists \delta < 1 \forall \delta \in (\delta, 1)$

\[ c^* - \xi \leq \liminf_{t \to \infty} a_{j,t} \leq \limsup_{t \to \infty} a_{j,t} \leq c^* + \xi \quad \forall j \in I^c = \{n + 1, \ldots, 2n\} \]

with probability 1.

Proof. Take $\xi > 0$ as given. Take $\zeta > 0$ to be adjusted. Since $V$ is strictly convex and compact, there exist $\zeta'$, $\zeta'' > 0$ such that

\[ \zeta' = \sup \{ \tilde{\zeta} \mid (r^* - \zeta, c^* + \tilde{\zeta}) \in V \}, \]

and

\[ \zeta'' = \sup \{ \tilde{\zeta} \mid (r^* - \zeta, c^* - \tilde{\zeta}) \in V \}. \]

See Figure 4. From (H.32), for all $j \in I^c$, if $a_j > c^* + 2\zeta'$, then $a_j < -\zeta$, and if $a_j < c^* - 2\zeta''$, then $a_j > \zeta$. Thus,

\[ c^* - 2\zeta'' < a_j < c^* + 2\zeta' \]

for any sufficiently large $t$. Note that as $\zeta \to 0$, $\zeta', \zeta'' \to 0$. Choose $\zeta$ so as to satisfy $\max \{2\zeta'', 2\zeta' \} \leq \xi$, and we obtain the result.
Figure 4. Case of $\gamma^* = 1$

References


**Department of Economics, University of Illinois, 1206 S. 6th Street, Champaign, IL 61820 USA**

*E-mail address: inkoocho@uiuc.edu*

*URL: https://netfiles.uiuc.edu/inkoocho/www*

**Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan**

*E-mail address: amatsui@e.u-tokyo.ac.jp*

*URL: http://www.e.u-tokyo.ac.jp/~amatsui*