games
MDPI

Article

Evolution of Mindsight and Psychological Commitment among Strategically Interacting Agents

Dimitry Rtischev

Faculty of Economics, Gakushuin University, 1-5-1 Mejiro, Toshima-ku, Tokyo 171-8588, Japan; dimitry.rtischev@gakushuin.ac.jp; Tel.: +81-3-5992-3075; Fax: +81-3-5592-1007

Academic Editor: Roman Sheremeta
Received: 27 May 2016; Accepted: 9 September 2016; Published: 21 September 2016

Abstract: We study the evolution of strategic psychological capabilities in a population of interacting agents. Specifically, we consider agents which are either blind or with mindsight, and either transparent or opaque. An agent with mindsight can observe the psychological makeup of a transparent agent, i.e., its logic, emotions, commitments and other elements that determine how it chooses actions. A blind agent cannot observe and opaque agents cannot be observed. Our assumption that mindsight and transparency are costly and optional exposes a middle ground between standard game theory without mindsight and evolution of preferences theory with obligatory and costless mindsight. We show that the only evolutionarily stable monomorphic population is one in which all agents are blind, opaque, and act-rational. We find that mindsight, transparency, and rule-rational commitments may evolve, albeit only in a portion of the population that fluctuates in size over generations. We reexamine the Ultimatum and Trust games in light of our findings and demonstrate that an evolved population of agents can differ significantly from a population of simplistic payoff-maximizers in terms of psychological traits and economic outcomes.

Keywords: theory of mind; mindsight; evolution of preferences; psychological commitment; act-rationality; rule-rationality; ultimatum game; trust game

JEL Classification: C73; D83; D87

1. Introduction

Most economic game models do not explicitly inquire into the origins of their players. The tacit assumption is that players are creatures of the moment, created out of nothing solely to earn as much as possible from playing a given game and then disappear without consequences. By construction, such momentary agents are "act-rational" in the sense that each player (i) always chooses the action that maximizes his payoff and (ii) assumes that other players also always choose their actions to maximize their payoffs. In the case of the Ultimatum game to divide a resource between a proposer and a responder, a responder who obeys (i) accepts any offer of one cent or more and a proposer who obeys (i) and (ii) offers one cent. But experiments reveal that human subjects usually do not play this way: Most proposers make substantial offers and many responders refuse small offers [1]. Evidently, many human responders do not obey (i) and most human proposers not obey (i) and/or (ii).

Experimental economics has documented such gaps between theoretical equilibria and experimental play in a variety of games. However, considering the different origins of momentary agents and human subjects, the gap in how they play should not surprise. Even anonymous strangers who interact only once in a carefully staged experiment are a product of a long process of evolution. According to evolutionary psychology, humans evolved under selection pressure favoring the ability to make psychological commitments and perceive or infer psychological commitments of others [2]. In particular, evolutionary psychologists stress that human interactions are fundamentally mediated by theory of mind, by subjective commitments secured with emotions not under voluntary control, and by other psychological capabilities refined through selection in social contexts [2]. At the level of the brain, interpersonal neurobiology emphasizes the role of "mindsight," which in that literature refers to the ability to dynamically perceive or infer the thoughts, feelings, beliefs, attitudes and/or intentions of another person when interacting with him or her [3].

A branch of game theory that has gone beyond momentary agents in the direction pointed by evolutionary psychology is evolution of preferences theory, also known as the indirect evolutionary approach [4]. The theory shows that evolutionary selection on the basis of relative performance in a strategic interaction yields agents who are committed to pursue "subjective" payoffs different from the "objective" payoffs they actually earn in the interaction [5,6]. Asking "What to maximize if you must?" in a generic game, Heifetz et al. (2007) [7] formally demonstrated that strategic interaction inherently generates the incentive to commit to maximize something other than the objective payoffs, and that such commitments do not disappear under evolutionary dynamics. In a similar vein, Aumann (2008) [8] informally argued that evolution favors "rule-rational" agents committed to an optimal rule of behavior over "act-rational" agents who optimize one act at a time.

All this ultimately rests on Schelling's (1960) [9] insight that if players can make commitments they often find it advantageous to do so, and the commitments can drastically change how the game is played. It is by constructing their agents as momentary beings incapable of making commitments to preferences that standard game models find nothing but act-rationality. It is by constructing their agents as evolved beings that could have made commitments to preferences that evolution of preferences models discover rule-rationality. However, this finding of universal rule-rationality among evolved agents critically depends on the assumption that the agents can costlessly and reliably commit to, display and observe preferences. Recognizing that this assumption is the Achilles' heel of evolution of preferences theory, efforts have been made to confirm that rule-rational equilibria are robust against small degradations in communication among agents about their preferences. Specifically, it has been shown that rule-rational equilibria do not abruptly unravel if costless communication about preferences becomes, for exogenous reasons, slightly noisy [7] or occasionally impossible [10].

However, a more telling test of whether rule-rationality can evolve is to endogenize the display and observation of preferences as costly capabilities that are themselves subject to selection.[1] Guth and Kliemt (1998, 2000) [11,12] have done this in the Trust game: They allowed first-movers to choose whether to purchase noisy information about the preferences of the second mover before deciding whether to trust him with an investment. We go further by considering a larger class of two-stage sequential-move games and endogenizing not only the observation of preferences by leaders but also the display of preferences by responders.[2] Specifically, we allow leaders to be blind or with mindsight and responders to be opaque or transparent. The key assumption that drives fitness tradeoffs in our model is that both mindsight and transparency are slightly costly.[3] Because of these costs, the rule-rational population, in which all agents are transparent and have mindsight, can be invaded by blind and opaque agents. Reversing the findings of the costless evolution of preferences models, we show that universal rule-rationality is not an equilibrium but universal act-rationality is. However, we also find that there exists a rest point surrounded by closed orbits along which rule-rational transparent agents and agents with mindsight may coexist in significant proportions. We apply our

---

1 Dekel et al. (2007) [10] call for this in their conclusion: " ... in the long-run, observability itself is subject to evolutionary forces. A more general model would incorporate the evolution of observability."

2 However, unlike Guth and Kliemt (1998, 2000) [11,12], we do not model noisy information.

3 We consider a transparent agent as one who has a mechanism to advertise his preferences and an opaque agent as one who lacks such a mechanism. From this perspective, we assume that transparency is costlier than opaqueness. However, in certain contexts transparency may be the cheaper, do-nothing alternative and opaqueness the costlier alternative that requires effort to conceal information. Since responders derive a strategic advantage from having their preferences taken into account by leaders, if showing preferences is less costly than concealing them then all responders in an evolved population would be transparent.

results to Ultimatum and Trust games to demonstrate that endogenously arising rule-rationality enabled by mindsight and transparency may have a significant impact on economic performance in a population of evolved agents.

The distinction between rule-rationality and act-rationality has been long made by philosophers of rationality and morality, albeit using different terminology [13] (Chapter VI). Of particular relevance is Danielson's (1990) [14] pioneering book, which explicitly considers mindsight and transparency among strategically interacting agents. Danielson brings great clarity to the problem by conceiving agents as programs that may take as input and examine other agents' programs, and may allow themselves to be taken as input and examined by others' programs. Following Danielson, we stipulate that each agent embodies a "decision logic" that chooses his actions. Specifically, in our framework the decision logic of certain leaders is capable of examining the decision logic of certain responders, and the population shares of such mindsighted leaders and transparent responders are determined by endogenous evolutionary dynamics. Broadly speaking, the decision logic of an agent in our formulation is a form of subjective preferences possessed by agents in other evolution of preferences models.

Modeling agents as embodying decision logics rather than maximizing subjective payoffs helps brings to the fore and clearly separate two key factors: (1) thinking, emotions, calculation and other *methods* individual agents use to choose their actions and (2) *communication* about these methods between agents. Although useful as a theoretical device, the decision logic formulation is obviously far from actual human interactions. Rather than directly displaying and observing their mental make-up, humans rely on words, tone of voice, gaze, facial expressions, gestures, and other signals that carry (potentially deceptive) information about the state of their minds. To relate our formulation to actual human interactions it helps to consider a recent experiment in which "trustees" in a Trust game could send a text message prior to playing a second and final round of the game with the same anonymous and unseen "investor" [15]. Some trustees used the message to relate their feelings in the form of apology or regret about the first round and indicate their intentions for the second round. On average, sending such a message enabled trustees to elicit more trust from investors. The fact that many senders used messages to reveal their feelings and intentions, and that many receivers reacted to such "cheap talk" even though the experiment was an anonymous one-shot interaction suggests that humans have a predilection to proactively exercise what loosely corresponds to transparency and mindsight in our abstract framework. The result that some of the participants chose to send a blank message and that receivers reacted differently to blank messages loosely corresponds to opaqueness in our model. Whether what appears as cheap talk actually incurs psychological costs and the distortions due to deception remain open issues.

The rest of the paper is organized as follows. The next section lays out the formal framework. Section 3 presents equilibrium analysis. Section 4 applies the findings to explore the economic consequences in the context of the Ultimatum Game and the Trust Game. Section 5 concludes.

2. The Model

There are two large, separate populations of agents: "leaders" and "responders." A dyad is formed by randomly drawing one agent from each population. Each dyad plays a two-stage "base game" as follows: first the leader observes the responder and takes action $x \in X$ and then the responder observes x and takes action $y \in Y$, where X and Y are finite action sets of the leader and responder, respectively. The resulting payoffs are $\pi_1(x, y)$ to the leader and $\pi_2(x, y)$ to the responder.

We make the following definitions and assumptions concerning the cognitive structures and processes that agents use to choose actions.

Definition 1. *"Decision logic" is a deterministic function, algorithm, or program located within an agent that computes the action the agent takes.*

Definition 2. *"Theory of mind" is what the leader in a dyad believes to be the decision logic used by the follower in the dyad.*

Assumption 1. *Each responder uses a decision logic to compute his action. A responder's decision logic takes as input the leader's action.*

Assumption 2. *Each leader uses a decision logic to compute his action. A leader's decision logic takes as input a theory of mind.*

Since decision logics are deterministic, the base game is played in pure strategies. We denote a responder's decision logic by $\Theta : X \to Y$ and his action by $y = \Theta(x)$. We denote a leader's decision logic by $\Lambda : \Omega \to X$ and his action by $x = \Lambda(\Phi)$, where $\Phi : X \to Y$ is the theory of mind and Ω is the set of all possible decision logics that a responder may have (and therefore Ω is also the set of all possible theories of mind that a leader may have).

We do not assume that responders' decision logic optimizes or that leaders hold the correct theory of mind; rather, we leave that to be determined endogenously by evolution. However, we do require that each leader optimize given the theory of mind he holds, per the following assumption[4].

Assumption 3. *Each leader's decision logic maximizes his payoff given his theory of mind:*

$$\Lambda(\Phi) = \underset{x}{\operatorname{argmax}} \pi_1(x, \Phi(x)) \tag{1}$$

Every leader is either blind (B) or with mindsight (M) and every responder is either transparent (T) or opaque (O), in the sense of the following definitions.

Definition 3. *A "leader with mindsight" can distinguish between a transparent and an opaque responder, observe the decision logic of a transparent responder, and use this information to construct a theory of mind.*

Definition 4. *A "blind leader" cannot distinguish between transparent and opaque responders nor observe the decision logic of any responder. A blind leader constructs a theory of mind without any information about the responder in his dyad.*

Definition 5. *A "transparent responder" has a decision logic which can be observed by any leader with mindsight.*

Definition 6. *An "opaque responder" has a decision logic which cannot be observed by any leader.*

In terms of the above "psychological traits," there are four kinds of dyads: MT, MO, BT, and BO. Only in the MT dyad does the leader observe the responder's decision logic and can use it as the theory of mind.[5] In the MO, BT, and BO dyads, the leader cannot observe responder's decision logic and therefore necessarily relies on a hypothesized theory of mind, which may or may not be correct.

Many types of opaque and transparent responders may exist, differing in terms of their decision logics. The type of responder is specified by (θ, Θ), where $\theta \in \{T, O\}$ indicates transparency or opaqueness and Θ is the decision logic. The state of responder population is given by the population share vector[6] $\mathbf{q} = (q_1, \ldots, q_O, q_{O+1}, \ldots, q_{O+T})$, where $q_i \in [0, 1]$ is the share of the ith type of responder (θ_i, Θ_i), O is the number of opaque responder types, T is the number of transparent responder types, and $\sum q_i = 1$.

Many types of blind and mindsighted leaders may exist, differing in terms of their theory of mind. The type of leader is specified by (λ, Φ), where $\lambda \in \{B, M\}$ indicates blindness or mindsight and Φ is the theory of mind. The state of leader population is given by the population share vector

4 Assuming this makes the theory of mind and the presence or absence of mindsight be the locus of evolution. It is analogous to the assumption in evolution of preferences literature that agents maximize subjective payoffs, which are allowed to evolve endogenously.

5 In general, the leader in the MT dyad may ignore the decision logic that he observes and use another theory of mind. However, per Proposition 2, doing so cannot increase the leader's payoff.

6 Population share vectors $\mathbf{q}$ and $\mathbf{p}$ are column vectors.

$\mathbf{p} = (p_1, \ldots, p_B, p_{B+1}, \ldots, p_{B+M})$, where $p_i \in [0,1]$ is the share of the *i*th type of leader (λ_i, Φ_i), B is the number of blind leader types, M is the number of leader types with mindsight, and $\sum p_i = 1$.

The payoffs in the Blind-Opaque dyad are:

$$\Pi_{ij}^{BO} = \pi_1(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \quad (2)$$

to leader $i = 1, \ldots, B$ and:

$$\Pi_{ij}^{OB} = \pi_2(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \quad (3)$$

to responder $j = 1, \ldots, O$.

Displaying one's decision logic is a costly capability, and therefore a transparent responder incurs a cost $\tau > 0$ every time he plays. Mindsight is also a costly capability, and therefore a leader with mindsight incurs a cost $\mu > 0$ every time he plays. In the Mindsight-Transparent dyad the leader observes the responder's decision logic and the payoffs are:

$$\Pi_{ij}^{MT} = \pi_1(\Lambda(\Theta_j), \Theta_j(\Lambda(\Theta_j))) - \mu \quad (4)$$

to leader $i = B+1, \ldots, B+M$ and:

$$\Pi_{ij}^{TM} = \pi_2(\Lambda(\Theta_j), \Theta_j(\Lambda(\Theta_j))) - \tau \quad (5)$$

to responder $j = O+1, \ldots, O+T$.

In a Blind-Transparent dyad, even though the leader cannot see, the responder nevertheless incurs the cost of transparency[7], and the payoffs are:

$$\Pi_{ij}^{BT} = \pi_1(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \quad (6)$$

to leader $i = 1, \ldots, B$ and:

$$\Pi_{ij}^{TB} = \pi_2(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) - \tau \quad (7)$$

to responder $j = O+1, \ldots, O+T$.

In a Mindsight-Opaque dyad, even though the responder is opaque, the leader still incurs the cost of mindsight, and the payoffs are:

$$\Pi_{ij}^{MO} = \pi_1(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) - \mu \quad (8)$$

to leader $i = B+1, \ldots, B+M$ and:

$$\Pi_{ij}^{OM} = \pi_2(\Lambda(\Phi_i), \Theta_j(\Lambda(\Phi_i))) \quad (9)$$

to responder $j = 1, \ldots, O$.

We gather all of the above payoffs in two payoff matrices. The leaders' payoff matrix has $B+M$ rows and $O+T$ columns arranged as follows:

$$\mathbf{\Pi}_L = \begin{bmatrix} \mathbf{\Pi}^{BO} & \mathbf{\Pi}^{BT} \\ \mathbf{\Pi}^{MO} & \mathbf{\Pi}^{MT} \end{bmatrix} \quad (10)$$

where $\mathbf{\Pi}^{BO} = [\Pi_{ij}^{BO}]$ is the B-row O-column matrix of leader payoffs in blind-opaque dyads, $\mathbf{\Pi}^{BT}$ is the B-row T-column matrix of leader payoffs in blind-transparent dyads, $\mathbf{\Pi}^{MO}$ is the M-row O-column

[7] Transparency and mindsight are built-in mechanisms that are costly to maintain. We assume that agents who have such mechanisms incur their cost in all interactions, whether they use them or not.

matrix of leader payoffs in mindsight-opaque dyads, and $\mathbf{\Pi}^{MT}$ is the M-row T-column matrix of leader payoffs in mindsight-transparent dyads. Analogously, the responders' payoff matrix has *B*+*M* rows and *O*+*T* columns arranged as follows:

$$\mathbf{\Pi}_R = \begin{bmatrix} \mathbf{\Pi}^{OB} & \mathbf{\Pi}^{TB} \\ \mathbf{\Pi}^{OM} & \mathbf{\Pi}^{TM} \end{bmatrix} \tag{11}$$

The evolutionary dynamics occur as follows. During each generation many random dyads are formed to play the base game. Each type of leader (responder) accumulates fitness equal to the sum of the payoffs earned by that type of leader (responder) in the base game. At the end of a generation agents replicate and die. Replication occurs within the leader and responder populations separately.[8] Specifically, the expected fitness of each leader type given the state of the responder population is given by the expected fitness vector $\mathbf{V}^L \equiv \mathbf{\Pi}_L\mathbf{q}$. The population average fitness of leaders is $\overline{V}_L \equiv \mathbf{p}\cdot\mathbf{V}^L$. Analogously, the expected fitness of each responder type given the state of the leader population is $\mathbf{V}^R \equiv \hat{\mathbf{\Pi}}_R\mathbf{p}$, where $\hat{\mathbf{\Pi}}_R$ is the transpose of $\mathbf{\Pi}_R$. The population average fitness of responders is $\overline{V}_R \equiv \mathbf{q}\cdot\mathbf{V}^R$. The replicator dynamic for leaders and responders, respectively, is:

$$\dot{p}_i = p_i(V_i^L - \overline{V}_L), \quad i = 1, \ldots, B+M \tag{12}$$

$$\dot{q}_j = q_j(V_j^R - \overline{V}_R), \quad j = 1, \ldots, O+T \tag{13}$$

In the equilibrium analysis that follows, we will focus on rest points per the following definition.[9]

Definition 7. *A "rest point" is a population state of leaders and responders (**p**, **q**) that satisfies the following conditions for all leaders i = 1, . . . , B+M and responders j = 1, . . . , O+T:*

$$V_i^L = \overline{V}_L \quad if\ p_i > 0 \tag{14}$$

$$V_j^R = \overline{V}_R \quad if\ q_j > 0 \tag{15}$$

$$V_i^L \leq \overline{V}_L \quad if\ p_i = 0 \tag{16}$$

$$V_j^R \leq \overline{V}_R \quad if\ q_j = 0 \tag{17}$$

3. Equilibrium Analysis

Two decision logics will play a central role in our equilibrium analysis. Following Aumann (2008) [8], we will call them act-rationality and rule-rationality and define each as follows.

Definition 8. *An "act-rational responder" has decision logic*

$$A(x) = \underset{y}{\operatorname{argmax}}\,\pi_2(x, y) \tag{18}$$

8 We model evolution in two populations interacting by playing different roles in an asymmetric base game. As is commonly done, we could have introduced a stage in which role assignment is randomly determined before play in each dyad. This would have symmetrized the base game and made all players belong to a single population. Our approach allows us to maintain separate focus on mindsight and transparency.

9 The rest points we consider per definition 7 are "saturated" in the sense that if a type of agent is absent, it cannot invade with a positive rate of growth [16] (p. 36). We do not consider unsaturated rest points since they can be invaded by an absent type.

Definition 9. *A "rule-rational responder" has decision logic R(x) such that*

$$\nexists R' \text{ s.t. } \pi_2(\Lambda(R'), R'(\Lambda(R'))) > \pi_2(\Lambda(R), R(\Lambda(R))) \tag{19}$$

and

$$\forall x \neq \Lambda(R) \quad \pi_1(x, R(x)) < \pi_1(\Lambda(R), R(\Lambda(R))) \tag{20}$$

Decision logic A is the responder's payoff-maximizing reply to any action by any leader. Decision logic R is the responder's payoff-maximizing strategic commitment vis-à-vis a leader with mindsight. The first condition in Definition 9 ensures that no other commitment yields the responder a higher payoff when the leader heeds the responder's commitment and maximizes his own payoff accordingly. The second condition ensures that a leader who does something else gets a lower payoff.

We will denote the leader's payoff-maximizing strategy given the decision logic of the responder using the following shorthand:

$$x_A \equiv \Lambda(A) = \underset{x}{\mathrm{argmax}}\, \pi_1(x, A(x)) \tag{21}$$

$$x_R \equiv \Lambda(R) = \underset{x}{\mathrm{argmax}}\, \pi_1(x, R(x)) \tag{22}$$

We will also use the following notation to denote base game payoffs to leaders (*i* = *1*) and responders (*i* = 2), gross of the costs of mindsight and transparency:

$$\begin{aligned} \pi_i^{RR} &= \pi_i(x_R, R(x_R)) \\ \pi_i^{AA} &= \pi_i(x_A, A(x_A)) \\ \pi_i^{RA} &= \pi_i(x_R, A(x_R)) \\ \pi_i^{AR} &= \pi_i(x_A, R(x_A)) \end{aligned} \tag{23}$$

We confine attention to base games in which strategic commitment affects payoffs. This class of games is large and can be formally described as in Reference [7]. For our purposes, it suffices to assume the payoff structure of the base game is such that act-rationality and rule-rationality are distinct and the costs of mindsight and transparency are not too large relative to their effect on payoffs. Formally, these assumptions are as follows:

Assumption 4. *The base game payoffs satisfy the following*

(a) x_A *and* x_R *exist, are unique and distinct:* $x_A \neq x_R$
(b) *responder can increase its payoff by strategic commitment:* $\pi_2^{RR} > \pi_2^{AA}$
(c) *responder can weakly increase its payoff by reneging on strategic commitment after the leader's move:* $\pi_2^{RA} \geq \pi_2^{RR}$

Assumption 5. *For a responder facing a leader with mindsight, the cost of transparency is less than the benefit of strategic commitment:* $\tau < \pi_2^{RR} - \pi_2^{AA}$

Assumption 6. *For a leader facing a rule-rational transparent responder, the cost of mindsight is less than the benefit of heeding the responder's strategic commitment rather than assuming that the responder is act-rational:* $\mu < \pi_1^{RR} - \pi_1^{AR}$.

The following seven propositions are the results of our equilibrium analysis. Although the propositions center on act-rational and rule-rational agents, they do not assume that act- and rule-rationality are the only possible decision logics. The proofs in the Appendix A explicitly consider the possibility of invasion by mutants with other decision logics.

The first two propositions characterize the types of agents that may or may not be found in rest point populations.

Proposition 1. *At a rest point, every opaque responder is act-rational.*

Proof. All proofs are in the Appendix A.

Proposition 2. *At a rest point, if there are leaders with mindsight, then every such leader uses the transparent responder's decision logic as the theory of mind and believes that an opaque responder is act-rational. That is, a leader with mindsight in a dyad with a responder of type* (θ, Θ) *uses the theory of mind:*

$$\Phi^M = \begin{cases} \Theta \, if \, \theta = T \\ A \, if \, \theta = O \end{cases} \tag{24}$$

Next, we consider monomorphic populations and show that there is only one that is evolutionarily stable.

Proposition 3. *The only monomorphic population that is a rest point is: all responders are opaque and act-rational, all leaders are blind and hold act-rationality as the theory of mind. Moreover, this is population cannot be invaded by other types of leaders or responders.*

Proposition 4. *A population in which all leaders have mindsight is not a rest point.*

Proposition 5. *A population in which all responders are transparent and have the same decision logic is not a rest point.*

Finally, we consider polymorphic populations. We restrict our attention to bimorphic populations composed of act- and rule-rational agents which cannot be invaded by other types of agents. The next two propositions identify two rest points, one of which is unstable and the other neutrally stable in the sense of being surrounded by closed orbits.[10] Our analysis does not rule out the existence of polymorphic equilibria in which agents have decision logics and theories of mind other than act- and rule-rationality. The proofs make use of the following additional assumption on the payoff structure of the base game.

Proposition 6. *If* $\pi_2^{RA} > \pi_2^{AA}$ *and the cost of mindsight is sufficiently small, there exists a rest point at which a share* $m* \in (0,1)$ *of leaders have mindsight and assume that opaque responders are act-rational, all other leaders are blind and use rule-rationality as the theory of mind, a share* $t* \in (0,1)$ *of responders are transparent and rule-rational, and all other responders are opaque and act-rational. This rest point is given by:*

$$m* = \frac{\pi_2^{RR} - \pi_2^{RA} - \tau}{\pi_2^{AA} - \pi_2^{RA}} \tag{25}$$

$$t* = 1 - \frac{\mu}{\pi_1^{AA} - \pi_1^{RA}} \tag{26}$$

Moreover, (m, t*) is a neutrally stable rest point surrounded by closed orbits with the time frequencies of (m, t) along the orbits equal to (m*, t*).*

Proposition 7. *If the cost of mindsight is sufficiently small, there exists an unstable rest point at which some leaders have mindsight and assume that opaque responders are act-rational, all other leaders are blind and use act-rationality as theory of mind, some responders are transparent and rule-rational, and all other responders are opaque and act-rational.*

The foregoing propositions describe combinations of psychological traits and decision logics that may evolve among agents subject to selection based on their performance in a sequential dyadic

---

[10] In general, polymorphic rest points of an asymmetric evolutionary game are not asymptotically stable under the replicator dynamic ([17], Theorem 12.8).

interaction. The only monomorphic evolutionarily stable state is universal blindness, opaqueness, and act-rationality. Blind leaders and opaque responders cannot go extinct. Opaque responders are act-rational. Although neither a blind leader nor a leader with mindsight can see the decision logic of an opaque responder, the leader with mindsight has an advantage when paired with an opaque responder because by detecting opaqueness he can infer that the responder is act-rational. Although mindsight, transparency, and rule-rationality cannot be universal and cannot be present in asymptotically stable proportions, they may nevertheless be present in proportions that oscillate. In such populations, blind leaders assume that responders are rule-rational and leaders with mindsight assume that opaque responders are act-rational.

4. Application to Ultimatum and Trust Games

4.1. Evolution of Mindsight among Agents Playing the Ultimatum Game

To explore the economic implications of the above findings in a concrete setting, imagine an island populated by two species: pushers and pullers. Every minute, a random pusher finds a resource and a random puller appears nearby. Every resource has value e, a large positive integer. To extract the resource they must push and pull together. The pusher offers to give x to the puller after they extract the resource, where $x \in X \equiv \{0, 1, .., e\}$. The puller accepts or rejects the offer: $y \in Y \equiv \{accept, reject\}$. If the puller rejects, the agents go their separate ways and the resource rots away. If the puller accepts, the agents cooperate to extract and divide the resource. The pusher and puller payoff functions are, respectively:

$$\pi_1(x, y) = \begin{cases} e - x & if\, y = accept \\ 0 & if\, y = reject \end{cases} \tag{27}$$

$$\pi_2(x, y) = \begin{cases} x & if\, y = accept \\ 0 & if\, y = reject \end{cases} \tag{28}$$

At the end of each year, all pushers reproduce asexually and die. A pusher which has earned more leaves more offspring than a pusher which has earned less. Pullers also reproduce in the same way. We want to know: after many years, what kind of cognition, perception, and behavior will prevail on this island? We also want to consider economic efficiency by comparing the average per-dyad product realized on the island to the first-best baseline $P = e$, which is the product that would be realized if pushers and pullers could use costless institutions ensuring cooperation in every dyad.

Let a small positive integer ε be a decision logic parameter representing the minimum gain to an agent that makes it worth cooperating. The decision logic of an act-rational puller is then:

$$A(x) = \begin{cases} accept\ if\ x \geq \varepsilon \\ reject\ if\ x < \varepsilon \end{cases} \tag{29}$$

and the decision logic of a rule-rational puller is:

$$R(x) = \begin{cases} accept\ if\ x \geq e - \varepsilon \\ reject\ if\ x < e - \varepsilon \end{cases} \tag{30}$$

Pusher strategies are: $x_A = \varepsilon$ and $x_R = e - \varepsilon$. The payoffs under all the various combinations of decision logics are:

$$\begin{array}{ll} \pi_1^{AA} = e - \varepsilon & \pi_2^{AA} = \varepsilon \\ \pi_1^{RR} = \varepsilon & \pi_2^{RR} = e - \varepsilon \\ \pi_1^{AR} = 0 & \pi_2^{AR} = 0 \\ \pi_1^{RA} = \varepsilon & \pi_2^{RA} = e - \varepsilon \end{array}$$

By Proposition 3, the blind/opaque/act-rational population in which all pushers are (B, A) and all pullers are (O, A) is evolutionarily stable.

By Proposition 6, the following population is a rest point surrounded by closed orbits:

Pushers: (B, R) and (M, Φ^M) Population shares: $b = 1 - \frac{\tau}{e-2\varepsilon},\ m = \frac{\tau}{e-2\varepsilon}$

Pullers: (O, A) and (T, R) Population shares: $o = \frac{\mu}{e-2\varepsilon},\ t = 1 - \frac{\mu}{e-2\varepsilon}$

provided $\mu < \frac{\varepsilon(e-2\varepsilon)}{e-\varepsilon}$ and $\tau < e - 2\varepsilon$.

Table 1 presents a numerical example comparing economic performance in the monomorphic population and at the bimorphic rest point. In the monomorphic equilibrium all pushers offer the minimum and pullers always accept. There is no mindsight among pushers or transparency among pullers. All pullers are act-rational and all pushers believe that all pullers are act-rational. Mindsight, transparency and rule-rationality exist along closed orbits around the bimorphic rest point. In these populations too offers are never rejected since blind pushers believe that pullers are rule-rational and offer almost everything. The total product realized is only τ less than in the monomorphic equilibrium, but is allocated almost entirely to the pullers. Mindsight and transparency thus serve to reverse the allocation in favor of pullers.

Table 1. Economic performance in the Ultimatum game in the monomorphic equilibrium consisting solely of act-rational agents without mindsight or transparency, and at the bimorphic rest point where act- and rule-rational agents coexist.

		e = 100, ε = 5, μ = 2, τ = 1	
		Monomorphic	**Bimorphic**
Population			
Leaders (pushers)			
Blind (B, A)	a	1	0
Blind (B, R)	b	0	0.989
Mindsighted (M, Φ^M)	m	0	0.011
Responders (pullers)			
Opaque act-rational (O, A)	o	1	0.022
Transparent rule-rational (T, R)	t	0	0.978
Performance			
Leader average fitness	V_L	95	5
Responder average fitness	V_R	5	94
Total product realized	$P=V_L+V_R$	100	99
First-best product possible	e	100	100
Fraction of first-best realized	P/e	1	0.99
Leader share of product	V_L/P	0.95	0.05
Responder share of product	V_R/P	0.05	0.95
Fraction of dyads with rejected offers	a t	0	0

Without mindsight pushers exploit the act-rational pullers and this is a stable equilibrium. Since mindsight and transparency enable pullers to turn the tables and exploit the pushers, it can be said that pullers prefer to display their rule-rational decision logic but pushers prefer not to look. However, even though mindsight hurts them, pushers with mindsight can be present in an evolved population. Although mindsight hurts pushers, because it is locally adaptive it does not go extinct. As the numerical example shows, even a small fraction of pushers with mindsight may be enough to support transparency and rule-rationality among almost all pullers, and make the blind pushers adopt rule-rationality as their theory of the puller's mind.

4.2. Evolution of Mindsight among Agents Playing the Trust Game

To explore the economic implications in another concrete setting, imagine an island is populated by "getters" and "workers." Every minute a random getter finds a resource. Every resource has value *e*, a large positive integer. The getter can "invest" some portion $x \in X \equiv \{0, 1, \ldots, e\}$ of the resource to be worked on by a random worker who is nearby. The worker's effort multiplies the value of the investment by a factor of k, a positive integer bigger than or equal to 2. After finishing the work, the worker can give back any amount $y \in Y \equiv \{0, 1, \ldots, kx\}$ to the getter. The resulting payoffs are $\pi_1(x, y) = e - x + y$ to the getter and $\pi_2(x, y) = kx - y$ to the worker. At the end of each year, all getters reproduce asexually and die. A getter which has earned more leaves more offspring than a getter which has earned less. Workers also reproduce in the same way. We want to know: after many years, what kind of cognition, perception, and behavior will prevail on this island? We also want to consider economic efficiency by comparing the average per-dyad product realized on the island to the first-best baseline $P = ke$, which is the product that would be realized if getters and workers could use costless institutions safeguarding maximal investment in all dyads.

The decision logic of an act-rational worker never gives back anything to the getter: $A(x) = 0$.

Let a small positive integer ε be a decision logic parameter representing the minimum gain to an agent that makes it worth cooperating. The decision logic of a rule-rational worker minimally rewards those getters who invest everything and punishes all others:

$$R(x) = \begin{cases} e + \varepsilon \, if \; x = e \\ 0 \, if \; x < e \end{cases} \tag{31}$$

Getter strategies are: $x_A = 0$ and $x_R = e$. The payoffs under the various combinations of decision logics are:

$$\begin{array}{ll} \pi_1^{AA} = e & \pi_2^{AA} = 0 \\ \pi_1^{RR} = e + \varepsilon & \pi_2^{RR} = ke - e - \varepsilon \\ \pi_1^{AR} = e & \pi_2^{AR} = 0 \\ \pi_1^{RA} = 0 & \pi_2^{RA} = ke \end{array}$$

By Proposition 3, the blind /opaque/act-rational population in which all getters are (B, A) and all workers are (O, A) is evolutionarily stable.

By Proposition 6, the following population is a rest point surrounded by closed orbits:

Getters: (B, R) and (M, Φ^{M}) Population shares: $b = 1 - \frac{e+\varepsilon+\tau}{ke}$, $m = \frac{e+\varepsilon+\tau}{ke}$

Workers: (O, A) and (T, R) Population shares: $o = \frac{\mu}{e}$, $t = 1 - \frac{\mu}{e}$

provided $\mu < \frac{e\varepsilon}{e+\varepsilon}$ and $\tau < e(k-1) - \varepsilon$

Table 2 gives a numerical example comparing economic performance in the monomorphic population and at the bimorphic rest point. In the monomorphic equilibrium getters do not invest anything and the workers earn nothing. At the bimorphic rest point, blind getters believe that workers are committed to repay with interest and invest everything. Some of them are betrayed by act-rational workers. However, investment occurs in most dyads, the only exception being dyads in which a getter with mindsight is paired with an opaque act-rational worker. Mindsight and transparency serve to increase the average product but also allocate most of the gains to the workers. However, average fitness of both getters and workers is higher at the bimorphic rest point than in the monomorphic population.

This case shows that mindsight, transparency and rule-rationality can be critical for trust, can make all players better off on average, and may exist in evolved populations. Unlike in the Ultimatum Game, mindsight and transparency are incentive-compatible for all: It can be said that workers want to show their decision logic and getters want to see it. However, since mindsight is costly, a fraction of getters evolves to free-ride without mindsight. Such blind trusting getters in turn create a niche for opaque act-rational workers, who evolve to prey on them. However, as the numerical example in Table 2 shows, distrust, betrayal, opaqueness and act-rationality can all be very rare even if only a minority of the getters have mindsight.

Table 2. Economic performance in the Trust game in the monomorphic equilibrium consisting solely of act-rational agents without mindsight or transparency, and at the bimorphic rest point where act- and rule-rational agents coexist.

		$e = 100, \varepsilon = 5, \mu = 2, \tau = 1, k = 5$	
		Monomorphic	**Bimorphic**
Population			
Leaders (getters)			
Blind (B, A)	a	1	0
Blind (B, R)	b	0	0.788
Mindsighted (M, Φ^M)	m	0	0.212
Responders (workers)			
Opaque act-rational (O, A)	o	1	0.02
Transparent rule-rational (T, R)	t	0	0.98
Performance			
Leader average fitness	V_L	100	102.9
Responder average fitness	V_R	0	394
Total product realized	$P = V_L + V_R$	100	496.9
First-best product possible	k e	500	500
Fraction of first-best realized	P/(ke)	0.2	0.9938
Leader share of product	V_L/P	1	0.21
Responder share of product	V_R/P	0	0.79
Fraction of dyads with reciprocated trust	(b+m)t	0	0.98
Fraction of dyads with distrust	a + m o	1	0.00424
Fraction of dyads with betrayal	b o	0	0.01576

5. Conclusions

We studied the decision logics and the capabilities for showing and observing them which may evolve among randomly paired agents subject to selection based on their performance in a strategic interaction. We found that if mindsight and transparency are costly and optional, then the only evolutionarily stable monomorphic population consists of agents which are blind, opaque, and act-rational. We also showed that blind leaders and opaque responders cannot go extinct and opaque responders must be act-rational.

These results cast doubt on the robustness of the rule-rational preferences found in equilibria of evolution of preferences models that assume costless and obligatory mindsight and transparency (e.g., [7]). However, our analysis does not rule out the evolution of mindsight, transparency, and rule-rationality. Rather, we show that mindsight, transparency, and rule-rationality may be found in a significant but fluctuating share of the population. In such populations, blind leaders assume every responder is rule-rational and leaders with mindsight assume that opaque responders are act-rational. Thus, our analysis exposes a middle ground between the rule-rationality predicted by evolution of preferences models and act-rationality presumed by standard game models.

We applied our general findings to two specific contexts: Ultimatum Game and Trust Game. In both games, mindsight, transparency, and rule-rationality serve to allocate most of the surplus to responders. Given the zero-sum nature of the Ultimatum Game, mindsight does not engender new value and leaders are better off in the equilibrium without mindsight. However, in the Trust Game, both leaders and responders earn more in populations with mindsight and mindsight, transparency, and rule-rationality are essential for enabling the investment that generates the new value.

The generality of our findings is limited by the simplistic specification of how preferences are displayed and observed, namely that agents can either show/see the true preferences or not show/see any preferences at all. A more general model would include responders that can display any preferences—true, false, or none—and leaders with various powers of mindsight, some of which can detect false preferences or discern the true preferences. There is a connection here to the evolution of agents with varying levels of theory-of-mind sophistication that has been studied using "level-k" cognitive hierarchy models (see [18] and papers cited therein). Cognitive hierarchy models hold preferences constant but, consistent with our results, also find that evolved populations are heterogeneous in terms of theory-of-mind capabilities. The modest step we took in the direction of exploring the co-evolution of preferences and theory-of-mind abilities may be fruitfully followed up by a deliberate effort to integrate the cognitive hierarchy and evolution of preferences frameworks.[11]

More generally, the locus of traits related to displaying, observing, inferring and committing to preferences appears to constitute an important dimension along which strategically interacting agents can evolve. Further work on this dimension of evolution may help understand the complex psychology and behavior being revealed by experimental economics.

Acknowledgments: The author gratefully acknowledges comments by *Games* editors and anonymous reviewers, by seminar attendees at Tsukuba University, Waseda University, and Kyoto University, and by participants of the 7th Conference on Logic, Game Theory and Social Choice in Bucharest, the 8th Spain-Italy-Netherlands Meeting on Game Theory in Budapest, and the Economics and Biology of Contests Conference in Brisbane.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

The proofs below make use of the following definition.

Definition A1. *Responder decision logics* Θ *and* Θ' *are "different" (denoted by* $\Theta \neq \Theta'$*) if the corresponding responder payoff matrices are different:* $\mathbf{\Pi}_R \neq \mathbf{\Pi}'_R$*. Theories of mind* Φ *and* Φ' *are "different" (denoted by* $\Phi \neq \Phi'$*) if the corresponding leader payoff matrices are different:* $\mathbf{\Pi}_L \neq \mathbf{\Pi}'_L$*.*

Proof of Proposition 1. Suppose that in the population of responders there are some opaque agents with act-rational decision logic A and some opaque agents with a different decision logic $Z \neq A$. Since a leader facing an opaque responder in a dyad cannot see whether the responder has A or Z, the leader's action is the same against either type of responder. By Definition 8, replying to the leader's action using A yields the responder a higher payoff than replying using Z. Thus opaque Z-responders earn lower average fitness than opaque A-responders. Therefore the assumed population is not a rest point. Next suppose that all opaque responders are act-rational. A mutant opaque responder with decision logic Z has a lower fitness and therefore cannot invade.□

Proof of Proposition 2. When a leader with mindsight is paired with an opaque responder, he sees that the responder is opaque and uses some theory of mind. Suppose some leaders use A and some leaders use $Z \neq A$ as the theory of mind for opaque responders. Since, at a rest point all opaque responders are act-rational (Proposition 1), the leader who uses A earns a higher payoff than the leader who uses Z. Thus the two types of leader cannot coexist at a rest point. Furthermore, if all leaders with mindsight use A for opaque responders, a mutant leader using Z for opaque responders cannot invade.

When a leader with mindsight meets a transparent responder with decision logic Θ, he can use it as the theory of mind or may use a different theory $Z \neq \Theta$. By Assumption 3, using Z yields the leader a lower payoff, and thus lower fitness, than using Θ. Thus the two types of leader cannot coexist at a rest point. Furthermore, if all leaders with mindsight use Θ for transparent responders, a mutant leader using Z for transparent responders cannot invade.

Proof of Proposition 3. Suppose all leaders are of type (B, A) and all responders are of type (O, A). Since all responders are opaque, a mutant leader with mindsight can get no information but would incur the cost of mindsight. Since all responders are act-rational, a mutant leader with a different theory of mind would earn less fitness. Thus mutants with mindsight or different theory of mind cannot

[11] One such recent effort studies a population consisting of sophisticated agents with a slightly costly "theory of preferences" capability and naïve agents lacking such capability; the study shows that when there is continuous novelty, the sophisticated agents learn opponents' preferences and this can help them evolve to take over the entire population [19].

invade the leader population. Since all leaders are blind, a mutant responder who is transparent cannot influence any leader's action, but would incur the cost of transparency. A mutant responder who is not act-rational would earn less fitness than an act-rational responder. Thus mutant responders who are transparent or have a different decision logic cannot invade the responder population. That other possible monomorphic populations are not rest points is proved as Propositions 4 and 5. □

Proof of Proposition 4. Suppose all leaders have mindsight. By Proposition 2, if this is a rest point then all leaders have theory of mind Φ^M. From the first condition in Definition 9 it follows that, when playing against a population of (M, Φ^M) leaders, a (T, R) responder would earn more than a (T, Z) responder for any $Z \neq R$. Assumption 5 implies that a (T, R) responder would earn more than a (O, A) responder. By Proposition 1, an (O, A) responder would earn more than a (O, Z) responder for any $Z \neq A$. Thus, if all leaders have mindsight and the population is a rest point, then all responders must be (T, R). A mutant (B, R) leader would earn the same payoff as a (M, Φ^M) leader but save the cost of mindsight. Thus a mutant blind leader can invade the leader population. □

Proof of Proposition 5. Suppose there is a rest point at which all responders are of type (T, Θ). A blind (B, Θ) leader would play the same as a (M, Φ^M) leader, but save the cost of mindsight. But if the leader population evolves to consist of entirely (B, Θ) agents, then a mutant (O, A) responder can invade because it would earn more than a (T, Θ) responder.□

Proof of Proposition 6. Consider a population of leaders consisting of (B, R) and (M, Φ^M) types and a population of responders consisting of (O, A) and (T, R) types. The system belongs to the class of asymmetric evolutionary games analyzed by Gintis (2009) [17] (Section 12.17). In this proof and next, we follow his approach to solve for rests points and characterize their stability.

Let b be the share of blind leaders, m the share of leaders with mindsight, o the share of opaque responders, and t the share of transparent responders. The leader population is given by $\mathbf{p} = (b, m)$, where $b + m = 1$ and $b, m \in [0, 1]$. The responder population is given by $\mathbf{q} = (o, t)$, where $o + t = 1$ and $o, t \in [0, 1]$. The payoff matrices are:

Leaders:

$$\mathbf{\Pi}_L = \begin{bmatrix} \pi_1^{RA} & \pi_1^{RR} \\ \pi_1^{AA} - \mu & \pi_1^{RR} - \mu \end{bmatrix}$$

Responders:

$$\mathbf{\Pi}_R = \begin{bmatrix} \pi_2^{RA} & \pi_2^{RR} - \tau \\ \pi_2^{AA} & \pi_2^{RR} - \tau \end{bmatrix}$$

Adding a constant to each entry in a column of $\mathbf{\Pi}_L$ or in a row of $\mathbf{\Pi}_R$ does not affect the replicator dynamics. Thus we can simplify the payoff matrices as follows:

$$\mathbf{\Pi}_L{}' = \begin{bmatrix} 0 & \mu \\ \pi_1^{AA} - \pi_1^{RA} - \mu & 0 \end{bmatrix}$$

$$\mathbf{\Pi}_R{}' = \begin{bmatrix} 0 & \pi_2^{RR} - \pi_2^{RA} - \tau \\ \pi_2^{AA} - \pi_2^{RR} + \tau & 0 \end{bmatrix}$$

Using the population share of blind leaders b and population share of opaque responders o as state variables, we can express the replicator equations of the two populations as follows:

$$\dot{b} = b(1-b)(\alpha - \gamma o)$$
$$\dot{o} = o(1-o)(\beta - \delta b)$$

where

$$\begin{aligned}\alpha &= \mu > 0 \\ \beta &= \pi_2^{AA} - \pi_2^{RR} + \tau < 0 \\ \gamma &= \pi_1^{AA} - \pi_1^{RA} \\ \delta &= \pi_2^{AA} - \pi_2^{RA}\end{aligned}$$

The rest point is given by:

$$b* = \frac{\beta}{\delta} = \frac{\pi_2^{AA} - \pi_2^{RR} + \tau}{\pi_2^{AA} - \pi_2^{RA}},\ o* = \frac{\alpha}{\gamma} = \frac{\mu}{\pi_1^{AA} - \pi_1^{RA}}.$$

Since α and β have opposite signs, the rest point is neutrally stable, surrounded by trajectories which are closed orbits such that the time frequencies of *(b, o)* along the orbits equal *(b*, o*)* ([17], Theorem 12.9).

The share of mindsighted leaders and transparent responders are, respectively

$$m* = 1 - b* = \frac{\pi_2^{RR} - \pi_2^{RA} - \tau}{\pi_2^{AA} - \pi_2^{RA}},\ t* = 1 - o* = 1 - \frac{\mu}{\pi_1^{AA} - \pi_1^{RA}}$$

Since $\pi_2^{RA} > \pi_2^{AA}$ and $\pi_2^{RA} > \pi_2^{RR}$, this is an interior rest point.

Next, we need to establish that a mutant responder cannot invade the bimorphic population consisting of (O, A) and (T, R) responders. Since both types of leader play the same action against a mutant (O, Z≠A) as against an incumbent (O, A), by not playing the maximizing response the mutant earns strictly less than the incumbent (O, A). Since both types of leader play the same action against a mutant (T, A) as against an incumbent (O, A), the mutant earns τ less fitness than the incumbent (O, A). Lastly, consider a mutant responder of type (T, Z) such that Z≠R and Z≠A. Since R is the decision logic that induces a leader with mindsight to take the action which maximizes the responder's best-reply payoff, the mutant earns less than the incumbent (T, R) responder earns against a (B, R) or (M, Φ^{M}) leader.

Finally, we need to establish that a mutant leader cannot invade the bimorphic population consisting of (B, R) and (M, Φ^{M}) leaders. As was shown in the proof of Proposition 2, a leader with mindsight but with a theory of mind different from Φ^{M} earns less than a (M, Φ^{M}) leader. So we only need to consider a mutant blind leader with decision logic Z≠R. The expected fitness of such (B, Z) leader is $V_{BZ} = o\pi_1^{ZA} + t\pi_1^{ZR}$. The expected fitness of incumbent (M, Φ^{M}) leader is $V_M = o\pi_1^{AA} + t\pi_1^{RR} - \mu$. The mutant cannot invade if $V_{BZ} < V_M$, which reduces to

$$\mu < o(\pi_1^{AA} - \pi_1^{ZA}) + t(\pi_1^{RR} - \pi_1^{ZR}).$$

From Definition 8 it follows that $\pi_1^{AA} > \pi_1^{ZA}$. From the second condition in Definition 9 it follows that $\pi_1^{RR} > \pi_1^{ZR}$. Therefore, the upper bound on μ is positive and for any Z there exists a sufficiently small positive μ such that a (B, Z) mutant earns less than incumbent leaders. □

Proof of Proposition 7. The proof is analogous to the proof of Proposition 6. The difference lies in the fitness earned by blind leaders and by responders paired with blind leaders. Consider a population of leaders consisting of (B, A) and (M, Φ^{M}) types and a population of responders consisting of (O, A) and (T, R) types. The key parameters of the replicator dynamic are:

$$\begin{aligned}\alpha &= \pi_1^{AR} - \pi_1^{RR} + \mu < 0 \\ \beta &= \pi_2^{AA} - \pi_2^{RR} + \tau < 0 \\ \gamma &= \pi_1^{AR} - \pi_1^{RR} \\ \delta &= \pi_2^{AR} - \pi_2^{RR}\end{aligned}$$

The rest point is given by:

$$b* = \frac{\beta}{\delta} = \frac{\pi_2^{RR} - \pi_2^{AA} - \tau}{\pi_2^{RR} - \pi_2^{AR}},\ o* = \frac{\alpha}{\gamma} = \frac{\pi_1^{RR} - \pi_1^{AR} - \mu}{\pi_1^{RR} - \pi_1^{AR}}$$

Since α and β have the same sign, *(b*, o*)* is a saddle point and therefore unstable ([17], Theorem 12.9).

References

1. Oosterbeek, H.; Sloof, R.; van de Kuilen, G. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Exp. Econ.* **2004**, *7*, 171–188. [CrossRef]
2. Nesse, R. *Evolution and the Capacity for Commitment*; Russell Sage Foundation: New York, NY, USA, 2001.
3. Siegel, D. Toward an interpersonal neurobiology of the developing mind: Attachment relationships, 'mindsight,' and neural integration. *Infant Ment. Health J.* **2001**, *22*, 67–94. [CrossRef]
4. Guth, W.; Yaari, M. An evolutionary approach to explain reciprocal behavior in a simple strategic game. In *Explaining Process and Change—Approaches to Evolutionary Economics*; Witt, U., Ed.; University of Michigan Press: Ann Arbor, MI, USA, 1992.
5. Banerjee, A.; Weibull, J. Evolutionary selection and rational behavior. In *Learning and Rationality in Economics*; Kirman, A., Salmon, M., Eds.; Blackwell: Oxford, UK, 1995; pp. 343–363.
6. Guth, W.; Peleg, B. When will payoff maximization survive? An indirect evolutionary analysis. *J. Evolut. Econ.* **2001**, *11*, 479–499.
7. Heifetz, A.; Shannon, C.; Spiegel, Y. What to maximize if you must. *J. Econ. Theory* **2007**, *133*, 31–57. [CrossRef]
8. Aumann, R. *Rule-Rationality versus Act-Rationality*; Discussion Paper 497; Center for Rationality and Interactive Decision Theory, Hebrew University: Jerusalem, Israel, 2008.
9. Schelling, T. *The Strategy of Conflict*; Harvard University Press: Cambridge, UK, 1960.
10. Dekel, E.; Ely, J.; Yilankaya, O. Evolution of Preferences. *Rev. Econ. Stud.* **2007**, *74*, 685–704. [CrossRef]
11. Guth, W.; Kliemt, H. The indirect evolutionary approach: bridging the gap between rationality and adaptation. *Ration. Soc.* **1998**, *10*, 377–399. [CrossRef]
12. Guth, W.; Kliemt, H. Evolutionarily stable co-operative commitments. *Theory Decis.* **2000**, *49*, 197–221. [CrossRef]
13. Gauthier, D. *Morals by Agreement*; Oxford University Press: Oxford, UK, 1986.
14. Danielson, P. *Artificial Morality: Virtuous Robots for Virtual Games*; Routledge: London, UK, 1992.
15. Schniter, E.; Sheremeta, R.M.; Sznycer, D. Building and rebuilding trust with promises and apologies. *J. Econ. Behav. Organ.* **2013**, *94*, 242–256. [CrossRef]
16. Sigmund, K. *The Calculus of Selfishness*; Princeton University Press: Princeton, NJ, USA, 2010.
17. Gintis, H. *Game Theory Evolving*, 2nd ed.; Princeton University Press: Princeton, NJ, USA, 2009.
18. Mohlin, E. Evolution of theories of mind. *Games Econ. Behav.* **2012**, *75*, 299–318. [CrossRef]
19. Robalino, N.; Robson, A. The Evolution of Strategic Sophistication. *Am. Econ. Rev.* **2016**, *106*, 1046–1072. [CrossRef]