Master in Economics: Empirical Applications and Policies

University of the Basque Country - UPV/EHU

Academic Course 2019/20

Master's Thesis

Common p-Belief in Rationality in Centipede Games

Aleix Garcia Galocha
Advisors: Jaromír Kovářík and Peio Zuazo-Garin

Common p-Belief in Rationality in Centipede Games*

Aleix Garcia-Galocha

Advisors: Jaromír Kovářík   Peio Zuazo-Garin

July 30, 2020

Abstract

Centipede Games represent a classic puzzle in game theory. In this work, we employ p-beliefs to show that almost any behavior is consistent with rationality and *almost* Common Belief in Rationality. However, Common p-Belief in Rationality cannot justify why people cooperate in some Centipede Games but not others in a non-trivial way. We thus propose a novel theoretical framework that links the p-beliefs in rationality to the incentives to cooperate. This more general subjective belief-based approach serves as a predictor of cooperation. We show that the proposed approach organizes well the behavior in an experiment with a large variety of Centipede Games.

Keywords. p-beliefs, Rationalizability, Centipede Games, Strategic Uncertainty

p-beliefs, rationalizability, Centipede Games, strategic uncertainty

*This paper corresponds to the Master's Thesis of the Master in Economics: Empirical Applications and Policies at the University of the Basque Country (UPV/EHU). I would like to thank the support, guidance and advice of Jaromír Kovářík and Peio Zuazo-Garin. All mistakes are my own

1 Introduction

The Centipede Game (CG, henceforth), proposed by Rosenthal (1981), represents a classic puzzle in game theory. In such game, two players choose alternately between two possible actions, *Pass* or *Take*, for a known number of rounds. If any player plays *Take*, the game ends. CGs are characterized by a particular payoff structure: each player's payoff from playing *Take* in a decision node is (i) lower than the payoff she gets if she plays *Take* in any later decision node, but (ii) higher than when she plays *Pass* and her opponent plays *Take*.[1] *Rationality and Common (strong) Belief in Rationality* (RCBR, henceforward) implies that the only subgame perfect Nash equilibrium (SPNE, hereafter) is *Take* at every decision node, resulting in a unique prediction: the game is over at the very first decision node.[2] However, people frequently play *Pass* initially in CGs, a behavior that stands in stark contrast with the theoretical prediction (see McKelvey and Palfrey, 1992; Rapoport, Stein, Parco and Nicholas, 2003; Bornstein, Kugler and Ziegelmeyer, 2004). Due to the payoff structure, which simultaneously incentivizes passing and taking before the opponent does, a conflict between payoff maximization and sequential reasoning arises. This classical tension is also reflected in other strategic situations, such as the classic repeated Prisoner's dilemma (see Dal Bó and Fréchette, 2011; Friedman and Oprea, 2012; Bigoni, Casari, Skrzypacz and Spagnolo, 2015; Embrey, Fréchette and Yuksel, 2018 for recent research), industry oligopolies or public good provisions. So it is important to understand how people behave when facing this type of situations.

In this paper, we first show theoretically that the standard prediction of behavior in CGs is extremely non-robust to relaxations of *Common (strong) Belief in Rationality* (CBR, hereafter). To this aim, we introduce the notion of p-rationalizability allowing players to entail uncertainty about higher-order mutual rationality. The result establishes that cooperation in CGs is justified for any $p < 1$ that is, no matter how negligible the uncertainty about higher-order rationality is.

However, this result is insensitive on the payoff structure as long as the game is a CG, while it has been documented that subjects' behavior differs systematically across different CGs (Fey, McKelvey and Palfrey, 1996; Kawagoe and Takizawa, 2012; Garcia-Pola, Iriberri and Kovářík, 2020). Our second contribution is to propose a subjective belief-based approach that can organize the behavior across both subjects and CGs with different payoff structures. To this aim, we combine two concepts: players' subjective beliefs about opponent's behavior and Threshold Conjectures that reflect the incentives to *Take* or *Pass*

[1] Section 2.1.2 formally characterizes the Centipede Game and Figure 1 displays an example of a general CG.

[2] Rationality and common strong belief in rationality is the natural counterpart of rationality and common belief in rationality for dynamic games, see Battigalli and Siniscalchi (2003).

in a particular CG. In particular, Threshold Conjectures correspond to subjective beliefs regarding opponents' behavior that induce indifferences between all the possible strategies at every decision node. Our methodology predicts that the higher the Threshold Belief is, the higher observed rates of cooperation should be, at a particular decision node.

Using lab data from an experiment with a large variety of CGs , we corroborate this positive relationship between Threshold Beliefs and the degrees of cooperation estimating correlations and some logit models. We find that the Threshold Conjectures explain well the behavior in experimental data at both the population and the individual level. Specially, for 60.56% of the subjects the Threshold Conjectures explain significantly the observed behavior across their 16 decisions in the 16 variations of the CG.

We contribute to two streams of literature. Related to epistemic game theory we introduce the *strong p-belief* operator, which building on Monderer and Samet's (1987) notion of p-belief generalizes Battigalli and Siniscalchi's (2003) strong belief, by requiring that an event is believed with probability at least p (instead of 1) at every *unexpected* history the event is consistent with (instead at every history the event is consistent with). The above allows for a more purely game-theoretic contribution consisting on the introduction of a new solution concept for dynamic games, (extensive-form) p-rationalizability, which generalizes extensive-form rationalizability (Pearce, 1984 and Battigalli, 1997). p-Rationalizability captures the behavioral implications of rationality and common strong p-belief in rationality and hence, it allows for representing arbitrarily small departures from the benchmark of perfect rationality in dynamic games. When $p = 1$, p-rationalizability and extensive-form rationalizability coincide, and when the game is static, our solution concept coincides with Hu's 2007 notion of p-rationalizability for static games. Applying p-rationalizability we show the extreme non-robustness of the standard prediction, as presented above.

We contribute to the literature which study the mismatch between the observed and the predicted behavior in the CG by introducing a novel approach. This approach is able to explain the differences in the observed behavior at both the population and the individual level based on the strong p-belief operator. We can classify the explanations that state why observed behavior differs from the unique SPNE in the CG in broadly three categories. Models of bounded rationality, that assumes that individuals are not completely rational (see McKelvey and Palfrey, 1995; Fey, McKelvey and Palfrey, 1996; McKelvey and Palfrey, 1998), preference-based models, which lean on the assumption that players do not maximizes only their own payoff (e.g. McKelvey and Palfrey, 1992; Garcia-Pola, Iriberri and Kovářík, 2020) and models that relax the notion of CBR. A Bayesian equilibrium approach (see McKelvey and Palfrey, 1992) and level-k thinking model (see Kawagoe and Takizawa,2012; Ho and Su, 2013 for recent research) were proposed to relax this assumption. Although there is a wide variety of literature that analyse why the players differ from the SPNE prediction, in this paper, we focus on the relaxation of the CBR allowing players to *suspect* about the opponents' rationality.

This paper is structured as follows. Section 2 sets out the theoretical framework and presents our theoretical result apart from proposing the methodology used. Section 3 presents the data we use and the results that support our belief-based approach. Section 4 concludes and presents future research derived from our study. Additionally, Appendices A to E incorporate all the auxiliary material.

2 Theoretical Framework

This section introduces the theoretical framework. First, we formally present dynamic games as well as the concrete framework of the CGs. Then, we recall the concepts of *p-belief* and *p-Rationalizability* and extend them to extensive-form dynamic games.

2.1 Preliminaries

We describe first the general framework for a dynamic game and characterize the terms of *conjecture* and *sequential rationality*. Then, we formally define the CGs.

2.1.1 Dynamic Games

A *dynamic game* (with complete information) consists of a list $\Gamma = \langle I, (A_i)_{i\in I}, H, Z, (u_i)_{i\in I}\rangle$ where I is the finite set of *players* and:

- For each player i, A_i is a finite set of *actions*. A *history* represents the unfolding of the game and consists of finite sequence of possibly simultaneous choices, i.e. on a finite sequence of elements from $\{\emptyset\} \cup \bigcup_{J\subseteq I} A_J$, where $A_J := \prod_{i\in J} A_i$ for any $J \subseteq I$. We say that history h' *follows* history h, denoted by $h \preceq h'$, if h' obtains from adding finitely many possibly simultaneous choices to h.[3]

- H and Z are finite and disjoint sets of histories such that $(H \cup Z, \preceq)$ is a rooted and oriented tree with terminal nodes Z. Symbol $\emptyset$ denotes the ex ante stage of the game (i.e., the root of the tree) and histories in H and Z are referred to as *partial* and *terminal*, respectively. For any player i and partial history h, let $A_i(h)$ denote the set of actions *available* to i at h. Player i is *active* at h if $A_i(h)$ is nonempty; let H_i denote the set of these histories. We assume that: (i) a player is never the only active one twice in a row, and (ii) whenever a player is active, at least two actions are available to her.

[3] That is, when there exists some $(a^n)_{n\le N} \subseteq A$ such that $h' = (h; (a^n)_{n\le N})$.

In this context, the set of player i's *strategies* is $S_i := \prod_{h \in H_i} A_i(h)$ and, as usual, the set of strategy *profiles* is denoted by $S := \prod_{i \in I} S_i$ and the set of player i's opponents strategies, by $S_{-i} := \prod_{j \neq i} S_j$. Obviously, for each partial history h, each strategy s induces a unique terminal history $z(s|h)$. Let $S_i(h)$ and $S_{-i}(h)$ denote, respectively, the sets of player i and i's opponents strategies that reach partial history h, and $H_i(s_i)$, the set of player i's histories that can be reached when she chooses strategy s_i.[4] Finally, for each terminal history Z and for each player i, the *payoff function* is defined as $u_i : Z \to \mathbb{R}$.

Conjectures

A *conjecture* for player i formalizes player's beliefs about other players' behavior as the game unfolds. Formally, for each player i, it consists of a conditional probability system $\mu_i = (\mu_i(h))_{h \in H_i \cup \{\emptyset\}}$ such that: (i) for every history h, either initial or in which player i is active, $\mu_i(h)$ is a probability measure on $S_{-i}(h)$ that assigns probability 1 to $S_{-i}(h)$, and (ii) whenever possible, beliefs are updated following conditional probability. We say that history $h \neq \emptyset$ is *unexpected* for conjecture μ_i if $\mu_i(h')$ assigns null probability to reaching h at every history h' preceding h.[5]

Sequential rationality

Each strategy s_i and conjecture μ_i naturally induce a *conditional expected payoff* at each history $h \in H_i \cup \{\emptyset\}$:

$$U_i\left(\mu_i, s_i \left| h\right.\right) := \sum_{s_{-i} \in S_{-i}} \mu_i(h)[s_{-i}] \cdot u_i\left(z\left(s_{-i}; s_i \left| h\right.\right)\right).$$

A player is said to be *sequentially rational* when her strategy maximizes her conditional expected payoff at every history that it reaches. This is captured when player i chooses a strategy which is in the set of *best-replies* for conjecture μ_i, formally defined as:[6]

$$r_i\left(\mu_i\right) := \left\{ s_i \in S_i \left| s_i \in \bigcap_{h \in H_i(s_i)} \operatorname*{argmax}_{s_i' \in S_i} U_i\left(\mu_i, s_i' \left| h\right.\right) \right.\right\}.$$

[4] To be precise, $S_i(h) = \{s_i \in S_i | h \preceq z(s_{-i}; s_i|\emptyset) \text{ for some } s_{-i} \in S_{-i}\}$ and $S_{-i}(h) = \prod_{j \neq i} S_j(h)$ on the one hand, and $H_i(s_i) := \{h \in H_i | s_i \in S_i(h)\}$ on the other.

[5] That is, if $\mu_i(h)[S_{-i}(h')] = 0$ for every $h' \prec h$.

[6] Notice that, if we denote the set of all conjecture of player i as $\Delta^{H_i \cup \{\emptyset\}}(S_{-i})$, then $r_i : \Delta^{H_i \cup \{\emptyset\}}(S_{-i}) \rightrightarrows S_i$ is upper-hemicontinuous.

2.1.2 Centipede Games

In our analysis, we slightly simplify the notation above to taylor it to the specific structure of CGs.[7] The CG is a two-player extensive-form game with perfect information where players choose between their available actions alternately. Figure 1 displays an example of a generic version of such game. Formally, for a fixed integer $n \geq 2$ we represent a CG of depth n as a list $\Gamma_n := \langle A, H, Z, (u_i)_{i=1,2} \rangle$, where:

- $A := \{r, d\}$ is the set of actions available to each player at each of her turns: r (*right*) and d (*down*).[8]
- $H := \{(i, 1), \ldots, (i, n)\}$ is the set of *partial histories*, so that for each player i and each $k \leq n$, (i, k) is the k^{th} decision node of player i. In consequence, the histories in which player 1 is *active* are $H_1 := \{(i, k) | k = 1, \ldots, n\}$. Similarly, player 2's set of histories is $H_2 := \{(1, 1)\} \cup \{(2, k) | k = 1, ..., n\}$ with $(1, 1)$ being included as well to represent the fact that player 2 has beliefs about the game *before* any choice has been made.
- $Z := \{(h, d) | h \in H\} \cup \{(2, n, r)\}$ is the set of *terminal histories*, so that each (h, d) denotes the terminal history in which partial history $h = (i, k)$ has been reached and player i has played d, and $(2, n, r)$ denotes the terminal history in which both players have always chosen r.
- $(u_i)_{i \in I}$ has the following particular structure. First, at every history of hers, a player prefers: (i) playing down at a later node of herself than at the present one, and (ii) playing down immediately than playing right and see her opponent play down. Let denote $u_i(j, k, a)$ as the payoff for player i when the action a is played by player j at her decision node k. Formally, for each $k = 1 \ldots, n-1$ we have:[9]

$$u_1(1, k+1, d) > u_1(1, k, d) > u_1(2, k, d),$$

$$u_2(2, k+1, d) > u_2(2, k, d) > u_2(1, k+1, d).$$

  The above expressions characterize the tension inherent to CGs. On one hand, each player has incentive to proceed onward in the game, because of the payoff derived from choose d in her following decision node, $k+1$, is higher than in the current one, k. For example for player 1, it is reflected in $u_1(1, k+1, d) > u_1(1, k, d)$. On the

[7] Nevertheless, we maintain the previous general definitions to preserve the generality of the methodology in Section 2.2.

[8] In this section, we refer to these available actions as *right* and *down*, instead of *Pass and Take*, because they represent what we observe graphically.

[9] Typically it is also assumed that $u_2(2, 1, d) > u_2(1, 1, d)$; while potentially relevant for lab implementation, it is strategically irrelevant.

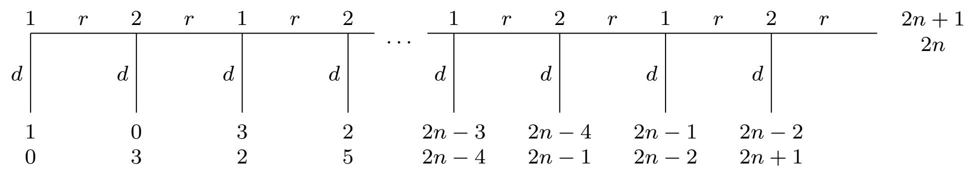


Figure 1: A Centipede Game of depth n

other hand, each player is better off if she chooses d before her opponent does. Again, for player 1 it is reflected in $u_1(1,k,d) > u_1(2,k,d)$. Finally, players have opposite preferences regarding the last and second to last terminal histories:

$$u_1(2,n,r) > u_1(2,n,d) \quad \text{and} \quad u_2(2,n,d) > u_2(2,n,r).$$

For convenience, we will make use of *reduced* strategies, so that the set of strategies of each player i can be conceived as:

$$S_i := \{s_i(k) | k = 1, \ldots, n, \infty\},$$

where, for every $k \leq n$, $s_i(k)$ consists of the strategy in which player i plays r until history (i,k), where d is played instead, and $s_i(\infty)$ consists of the strategy in which player i plays r at every history of hers. Reduced strategies to analyse the CG were firstly used by Nagel and Tang (1998), but are commonly applied in the current literature.

2.2 p-Rationalizability

A player is *rational* in a game if she chooses a strategy that maximizes the expected utility given a probabilistic belief over the opponents' behavior. The epistemic assumptions of RCBR – that is, everybody is rational, everybody (strongly) believes that everyone is rational, and so on – allows us to introduce the solution concept of (extensive-form) *Rationalizability* (Pearce, 1984 and Battigalli, 1997). We can relax the assumption of CBR by allowing players to entertain some suspicion that their opponents: (1) may not be rational (in any arbitrary sense), or (2) may suspect that their opponents are not rational, or (3) may suspect that their opponents suspect that their opponents are not rational, and so on. We formalize this idea of *suspicion* via Monderer and Samet's (1987) *p-belief*, which relaxes the notion of certainty by requiring that an event is believed with "at least probability p" (for some exogenously given p) instead of "with probability 1" as in case of rationality.

Under the mentioned relaxation of belief, we introduce a solution concept that generalizes Battigalli's (1997) formalization of Pearce's (1984) (extensive-form) rationalizability. We consider *Strong p-belief*, which is a relaxation of p-belief to dynamic settings. For each

$p \in [0,1]$, we say that an event $E_{-i} \subseteq S_{-i}$ is *strongly* p-believed by conjecture μ_i if two conditions hold. First, μ_i initially puts probability at least p on E_{-i}. Second, at every *unexpected* history h that might be reached by some strategy in E_{-i}, conjecture μ_i puts probability at least p on E_{-i}.[10] The p-belief constraint is not imposed on *every* history consistent with E_{-i}, but instead, among those consistent with E_{-i}, only on those that were assigned null probability at every preceding history. This allows for μ_i assigning very low probability to E_{-i} at some history E_{-i} is consistent with, but which was considered unlikely to be reached—yet not. Formally we get:

Definition 1 (Strong p-belief). *Let Γ be a dynamic game. Then, for every $p \in [0,1]$, every player i, every conjecture μ_i and every event $E_{-i} \subseteq S_{-i}$ we say that μ_i* strongly p-believes in E_{-i} *if $\mu_i(h)[E_{-i}] \geq p$ for every history h of either the following kind:*

(i) $h = \emptyset$.

(ii) *h is unexpected for μ_i and is reached by B_{-i}.*

Given this extension of p-belief, it is immediate to define the solution concept our analysis will rely on. This consists of an iterated elimination procedure that follows a very simple logic: for a fixed $p \in [0,1]$, in the first round we eliminate all strategies that are not a best-reply to some conjecture, in the second round we eliminate all strategies that are not a best-reply to some conjecture that strongly p-believes in the strategies of the opponents' that survived the first round, in the third round we eliminate all strategies that are not a best-reply to some conjecture that strongly p-believes in the strategies of the opponents' that survived the second round, and so on.[11] We can formally define, then, the solution concept in the following manner:

Definition 2 (p-Rationalizability). *Let $p \in [0,1]$. The set of* p-rationalizable *strategies of player i is defined as an iterated elimination process by setting $R^p_i := \bigcap_{k \geq 0} R^p_{i,k}$, where:*

$$R^p_{i,0} := S_i,$$

$$R^p_{i,k} := \left\{ s_i \in R^p_{i,k-1} \,\middle|\, s_i \text{ is a best-reply for some } \mu_i \text{ that } p\text{-believes in } R^p_{-i,k-1} \right\},$$

for every $k \in \mathbb{N}$.

[10] Remember that, in Section 2.1.1, we define that a history $h \neq \emptyset$ is *unexpected* for a conjecture, if, at every history that precedes, the conjecture assigns null probability to reaching h.

[11] Or, in other words, a player is rational, she strongly p-believes that her opponents are rational, she strongly p-believes that her opponents strongly p-believes that their opponents are rational, and so on. Technically, this assertion requires an epistemic characterization result, but this would follow from standard (yet tedious) arguments.

Remark 1. *Obviously, both notions monotonically become more stringent as p increases and, for $p = 1$, both converge to strong belief (see Battigalli and Siniscalchi, 2002) and (extensive-form) rationalizability, respectively. Clearly, in static settings Definition 2 coincides with Hu's (2007)* p-rationalizability.

After providing the theoretical result emanated from the general theoretical framework, Proposition 1 shows the above result applied to the CG.

Proposition 1. *Let Γ_n be a Centipede Game of depth n. Then, there exists some $\bar{p} \in (0,1)$ such that:*

$$R_1^p := \begin{cases} \{s_1(1)\} & \text{if } p = 1, \\ S_1 \setminus \{s_1(\infty)\} & \text{if } p \in (\bar{p}, 1), \\ S_1 & \text{if } p \in [0, \bar{p}], \end{cases} \qquad R_2^p := \begin{cases} \{s_2(1)\} & \text{if } p = 1, \\ S_2 \setminus \{s_2(\infty)\} & \text{if } p \in [0, 1). \end{cases}$$

Proof. Given the payoff structure of CG, the only strategy that is not sequentially rational in this game is $s_2(\infty)$. Hence we know that $R_{1,1}^p = S_1$ and $R_{2,1}^p = S_2 \setminus \{s_2(\infty)\}$ for every $p \in [0,1]$. We proceed now in two steps:

Preliminary observation. *There exists some $\bar{p} \in (0,1)$ such that $s_1(\infty) \in R_{2,1}^p$ for every $p \leq \bar{p}$.* Consider conjecture μ_1^0 that assigns probability 1 to $s_2(\infty)$ at every history of player 1's. Clearly, $s_1(\infty)$ is its unique sequential best reply and, furthermore, it is its unique conditional strict best reply at every history of player 1. Take arbitrary $p \in (0,1)$ and define conjecture μ_1^p as follows:

$$\mu_1^p(h) := \begin{cases} p \cdot 1_{\{s_2(1)\}} + (1-p) \cdot \mu_1^0(h) & \text{if } h = (1,1), \\ \mu_1^0(h) & \text{otherwise.} \end{cases}$$

Clearly, μ_1^p is a well-defined conjecture. Since $s_1(\infty)$ is a unique conditional best reply at every history of player 1, there exists some $\bar{p} > 0$ such that $s_1(\infty)$ is a sequential best reply to μ_1^p for every $p \leq \bar{p}$. Thus, since $R_{2,1}^p = S_2 \setminus \{s_2(\infty)\}$ for every $p \leq \bar{p}$, μ_1^p justifies the inclusion of $s_1(\infty)$ in $W_{1,2}^p$. The fact that we know that $s_1(\infty) \notin R_{1,2}^1$ lets us conclude that $\bar{p} \in (0,1)$. ★

Proof of the proposition. Fix arbitrary $p \in [0,1)$. The fact that $R_{1,1}^p = S_1$ implies that $R_{2,2}^p = R_{2,1}^p$. We know then that:

$$R_{1,2}^p := \begin{cases} S_1 \setminus \{s_1(\infty)\} & \text{if } p \in (\bar{p}, 1), \\ S_1 & \text{if } p \in [0, \bar{p}], \end{cases} \qquad R_{2,2}^p = S_2 \setminus \{s_2(\infty)\}.$$

We proceed now by induction: Suppose that the same equalities hold for some $m \geq 2$. We will show that they also to for $m+1$. For player 1 it is immediate: the fact that

$R^p_{2,m} = R^p_{2,m-1}$ implies that $R^p_{m+1} = R^p_m$. For player 2, the claim is trivially true of $R_{1,m} = S_1$. Otherwise, for each $k = 1, \ldots, n$ construct conjecture μ_2^p as follows:

$$\mu_2^{p,k}(h) := \begin{cases} p \cdot 1_{\{s_1(k)\}} + (1-p) \cdot 1_{\{s_1(\infty)\}} & \text{if } h = (2, k') \text{ with } k' < k, \\ 1_{\{s_1(\infty)\}} & \text{otherwise.} \end{cases}$$

Obviously $\mu_2^{p,k}$ is a conjecture that justifies the inclusion of $s_2(k)$ in $R^p_{2,m+1}$ and hence the proof is complete. ■

From Proposition 1, it follows that cooperation in CG is justified even by an epsilon doubt about the belief in rationality of the opponent. In fact, almost any behavior is p-rationalizable under the minimum suspicion. Moreover, *Pass Always* is also p-rationalizable for player 1 under a larger degree of doubt ($p \leq \bar{p}$). Actually, $\bar{p}$ is the probability that allows for player 1 being indifferent between playing *Pass* or *Take* in her last decision node. Obviously, under no suspicion, that is with $p = 1$, the only p-rationalizable strategy for both players is choose *Take* in their first node, as under the notion of rationalizability. Hence, we detect one element of the theoretical prediction that is extremely non-robust even to a tiny relaxation: an arbitrarily small uncertainty regarding the opponents' rationality rationalizes any degree of cooperation.

Second, player 2's behavior is trivially identified as it does not depend of $\bar{p}$ (except for the extreme case of p=1). Thus, we cannot use the data from player 2 for the identification purpose. This observation is in line with the one presented by Brandenburger, Danieli and Friedenberg (2019).

Note that Poposition 1 has an important limitation: p-rationalizability cannot discriminate behavior across different subjects and/or different CGs. We target this issue in the following section.

2.3 Threshold Conjectures

Proposition 1 shows that Common p-Belief in Rationality rationalizes any strategy except $s_i(\infty)$ for both players for any $p < 1$, and even legitimizes $s_1(\infty)$ for $p \leq \bar{p}$. Therefore, p-rationalizability does not allow to explain behavioral differences across subjects and across different variations of CGs. In this section, we propose a novel belief-based approach to overcome this issue. Our methodology allows us to explain different behavior across subjects and/or games on basis of differences in subjective beliefs about opponents' rationality and differing incentives to play *Take* or *Pass* across different CGs.

Let m denote the decision node perspective, $m \in \{1, \ldots, n,\}$, and k label the decision node of the opponent in which she plays *Take*, $k \in \{m, m+1, \ldots, n, \infty\}$. Then, we define the subjective (updated) belief $\mu_1^m(k) := \mu_1(m)[s_2(k)]$ as the belief of player 1 in decision node

m that the opponent plays $s_2(k)$. Consequently, we can define $\mu_1^m := (\mu_1^m(k))_{k=m,m+1,\dots,n,\infty}$ as the player 1's conjecture, at node m. We call subjective *updated* conjectures the belief formed in later nodes, i.e. when $m > 1$. For instance, consider a CG with $n = 3$. A generic conjecture of player 1 is denoted by $\mu_1^1 = \left(\mu_1^1(1), \mu_1^1(2), \mu_1^1(3), \mu_1^1(\infty)\right)$, where $\mu_1^1(1)$ represents the probability with which player 1 believes that her opponent chooses $s_2(1)$. A generic updated conjecture from the perspective of node 2 is denoted by $\mu_1^2 = \left(\mu_1^2(2), \mu_1^2(3), \mu_1^2(\infty)\right)$ and, finally, the subjective updated conjecture at node 3 as $\mu_1^3 = \left(\mu_1^3(3), \mu_1^3(\infty)\right)$. It is important to note that the subjective conjectures are not objective, material elements, but rather, subjective personal assessments of uncertainty of each individual. Importantly, individual differences in the subjective beliefs can explain why some players cooperate but not others in the same CG.

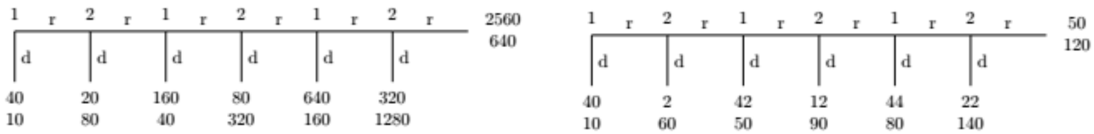
For each CG at each decision node of player 1, H_1, we obtain the *(Updated) Threshold Beliefs* $(\bar{\mu}_1^m(k))_{k=m,\dots,n,\infty}$, from node m perspective, that make player 1 indifferent between $s_1(k)$ and $s_1(k+1)$, whenever possible beliefs are updated following conditional probability. Thanks to the full support of the (Updated) Threshold Beliefs, we can also define the *(Updated) Threshold Conjecture* of a game, $\bar{\mu}_1^m := (\bar{\mu}_1^m(k))_{k=m,m+1,\dots,n,\infty}$, which is the conjecture that maximizes player 1 strategic uncertainty. That is, a subject in player 1's role with $\mu_1^1 = \bar{\mu}^1$ would be indifferent among *all* her strategies at every node of the game. Consider, again, a CG with $n = 3$. The prior Threshold Conjecture is defined as $\bar{\mu}_1^1 = \left(\bar{\mu}_1^1(1), \bar{\mu}_1^1(2), \bar{\mu}_1^1(3), \bar{\mu}_1^1(\infty)\right)$, where $\bar{\mu}_1^1(1)$ represents the belief that make player 1 indifferent between $s_1(1)$ and $s_1(2)$.[12] The Updated Threshold Conjecture from node 2 perspective is defined as $\bar{\mu}_1^2 = \left(\bar{\mu}_1^2(2), \bar{\mu}_1^2(3), \bar{\mu}_1^2(\infty)\right)$ and the Updated Threshold Conjecture at node 3 as $\bar{\mu}_1^3 = \left(\bar{\mu}_1^3(3), \bar{\mu}_1^3(\infty)\right)$.[13] While the subjective conjectures, which are intrinsic to the decision makers, explain the differences in the observed behavior across subjects playing the same CG, the Threshold Conjectures are computed from the game payoffs and can thus predict different behavior across CGs.

Remember that Player 2's p-rationalizable strategies do not depend on $\bar{p}$ as shown in Proposition 1. As a result, player 2 data cannot be used to check our theoretical results. For this reason, we only define both (Updated) Threshold Beliefs and (Updated) Threshold Conjectures for player 1 and, therefore, in Section 3 we focus our empirical analysis only on this player role. To simplify the notation, we will leave aside the subscript referring players role i and substitute it to a subscript that refers to the decision makers.

Now, we clarify the interpretation of our theoretical framework for a CG with $n = 3$ because the empirical data used in Section 3 presents this length. Consider a subject i

[12] Appendix A shows how we obtain the prior Threshold Conjecture for the McKelvey and Palfrey's (1992) *exponentially increasing sum* Centipede Game, with $n = 3$.

[13] Note that as the game unfolds the (Updated) Threshold Conjecture decreases by n, as in the case of subjective (updated) conjectures



(a) CG_A (b) CG_B

Figure 2: Different CGs with Different Prior Threshold Conjectures

who has a prior conjecture denoted by $\mu_i^1 = \left(\mu_i^1(1), \mu_i^1(2), \mu_i^1(3), \mu_i^1(\infty)\right)$. As mentioned above, each (Updated) Threshold Conjecture $\bar{\mu}^m$ represents the beliefs that make player 1 indifferent at node m between choose any of the available strategies. Focusing on the first decision node, if $\mu_i^1(1) \geq \bar{\mu}^1(1)$ *Take First* belongs to the set of best-replies for a given subjective belief, while for $\mu_i^1(1) < \bar{\mu}^1(1)$ *Take First* does not belong to the set of best-replies. That is, if the individual believes that her opponent would choose *Take* in her first node with a higher probability than the prior Threshold Belief of the CG, *Take First* belongs to the set of player 1's best-replies. Therefore, given a distribution of $\mu_i(1)$ obtained from a sample of different decision makers, different CGs with different $\bar{\mu}(1)$ must imply different theoretical predictions.

Example 2.1. *Consider that subject i with $\mu_i^1(1) = 0.75$. That is, this subject believes that her opponent play down in her first decision node, i.e. chooses Take First, with a probability equal to* 0.75*. There are two possibilities: (i) the subjective prior belief is lower than the prior Threshold Belief $\bar{\mu}^1(1)$ and (ii) the subjective prior belief is higher or equal than the prior Threshold Belief $\bar{\mu}^1(1)$. Consider the games shown in Figure 2, CG_A and CG_B, with $\bar{\mu}^1(1) = 0.8571$ and $\bar{\mu}^1(1) = 0.05$, respectively. Since $\mu_i^1(1) = 0.75 < \bar{\mu}^1(1) = 0.857$ in CG_A, the best response set contains all possible strategies except Take First while in the CG_B, where $\mu_i^1(1) > \bar{\mu}^1(1)$, Take First is the only best response. So, the theoretical predictions under our methodology are any strategy except Take First in the first CG and Take First in the second CG. In a pool of subjects, with different subjective beliefs, playing CG_A and CG_B we expect more people choosing Take First in the latter game than in the former. In other words, if this Threshold Belief is very high, many different players would have their subjective beliefs below it and Take First would not belongs to the best-replies sets of their observing, then, higher cooperation rates.*

The example above shows how different (Updated) Threshold Beliefs imply different theoretical predictions. In particular, they may predict why different people behave heterogeneously in the same CG and why same people act different across modified versions of the CG. CG_A and CG_B correspond to $CG1$ and $CG5$ of the data used in Section 3 with an observed *Take First* frequencies equal to 3.95% and 65.79%, respectively, corroborating the prediction of the proposed theory.

For players who do not *Take First*, the decision of *Take Second* or not would follow the same reasoning as in the Example 2.1. But, instead of analyse $\bar{\mu}^1(1)$ and her subjective belief $\mu_i^1(1)$, the individual must take into account the Updated Threshold Belief $\bar{\mu}^2(2)$ compared with her subjective updated belief $\mu_i^2(2)$. The same applies for the third, and last, decision node. Thus, contingent upon the structure of (Updated) Threshold Beliefs of every game the predicted strategies may vary across CGs with different payoffs structures. This is the relationship we explore empirically in the section below.

3 Empirical Analysis

In this section, we apply the proposed theoretical methodology to experimental data. We present the data and then analyse the relationship between Threshold Conjectures and observed behavior at both the population and the individual level.

3.1 Data

The data used come from Garcia-Pola, Iriberri and Kovářík (2020) where the authors propose 16 different CGs with $n = 3$, presented in Appendix D, with a large variety in the payoff structure. Authors classify the games according to some characteristics. First, half of the games start with unequal splits while the rest games have perfect equality in the very first node. Second, the evolution of the sum of players' payoffs in each terminal node can be differentiated in four groups: increasing-sum, constant-sum, decreasing-sum and variable-sum.[14] The participants in the experiment decide, using reduce-form strategies, in the 16 different CGs via strategic-method and without feedback between games to reduce as much as possible learning and reputation concerns. As we show above, the (Updated) Threshold Conjectures of a CG depend on the payoffs structures. Therefore, thanks to the payoff differences across the 16 CGs, we obtain a variety of (Updated) Threshold Conjectures' structures, which we exploit in the next section. Table 1a lists the corresponding prior Threshold Conjecture $\bar{\mu}^1$ of each CG and Figure 6 in Appendix E illustrates their evolution graphically.[15] Note that, accordingly with the evolution of prior Threshold Conjectures, there is no common trend among all games, but we can roughly classify them as increasing, decreasing, V-shaped and Threshold Conjectures with no specific evolution.

Table 1b displays the observed relative frequencies of each strategy played by decision makers in the role of player 1 (P1, henceforth) for each of the 16 CGs. We can highlight that the modal choices vary between groups of games. For increasing-sum games, the modal

[14]Games from 1 to 8 are classified as increasing-sum, games 9 and 10 as constant-sum, games 11 and 12 as decreasing-sum and the rest as variable-sum

[15]Table 6 in Appendix C displays the Updated Threshold Beliefs from node 2 and 3 perspective.

choices are more concentrated around *Take Second* and *Take Third*, except for CG 5 and 6 where the most frequent choice is *Take First* with a relative frequencies higher than 50%. While in decreasing-sum games the modal strategy is *Take First*. In the constant-sum games the most frequent election is *Take Second* and, finally, for variable-sum games there is no a common pattern but the *Take First* is the modal choice for almost all the games.

CG	$\bar{\mu}^1(1)$	$\bar{\mu}^1(2)$	$\bar{\mu}^1(3)$	$\bar{\mu}^1(\infty)$
1	85.71	12.24	1.75	0.29
2	93.18	6.01	0.73	0.08
3	80.00	10.00	3.64	6.36
4	90.91	4.46	1.08	3.55
5	5.00	5.94	19.08	69.98
6	11.76	16.54	32.26	39.43
7	96.15	3.21	0.53	0.11
8	65.91	11.36	18.18	4.55
9	63.64	7.06	2.65	26.65
10	66.67	15.00	1.47	16.87
11	33.33	9.52	5.19	51.95
12	37.50	5.21	6.03	51.26
13	12.50	40.00	15.83	31.67
14	28.57	31.43	13.75	26.25
15	14.29	7.79	3.32	74.61
16	75.00	1.32	0.95	22.74

(a) *Prior Threshold Conjectures*

CG	$s(1)$	$s(2)$	$s(3)$	$s(\infty)$
1	3.95	32.89	40.79	22.37
2	2.63	34.21	31.58	31.58
3	15.79	57.89	18.42	7.89
4	9.21	64.47	21.05	5.26
5	65.79	14.47	13.16	6.58
6	51.32	15.79	19.74	13.16
7	15.79	21.05	25.00	38.16
8	53.95	21.05	14.47	10.53
9	22.37	59.21	11.84	6.58
10	11.84	67.11	15.79	5.26
11	64.47	10.53	15.79	9.21
12	55.26	32.89	7.89	3.95
13	50.00	17.11	22.37	10.53
14	31.58	39.47	15.79	13.16
15	72.37	10.53	14.47	2.63
16	39.47	40.79	10.53	9.21

(b) Relative Frequencies

Table 1: Data across CGs, Player 1

3.2 Results

This section reviews, first, whether there is the predicted negative relationship between the Threshold Conjectures and the observed strategies. Then, we provide various evaluations of how our theoretical result can explain the observational data. To this purpose, we apply different techniques to check by how much our methodology explicate the observed behavior at the population and the individual level.

Table 2 shows the correlations between the dummy variables for each strategy and the corresponding potentially (Updated) Threshold Conjectures.[16] The diagonal elements of the table support the intuition presented in Section 2.3. At each decision node, the higher the Threshold Belief $-\bar{\mu}^1(1), \bar{\mu}^2(2),\ \bar{\mu}^3(3)-$ the lower the likelihood to choose *Take* and, consequently, the lower the observed frequencies of *Take* in that node. Moreover, although the magnitude of the estimated correlation coefficients decrease as the game unfolds, all the

[16] The correlations are computed using only the behavior of subjects who reach the corresponding node in the corresponding game for $m > 1$.

diagonal elements are significant at 1% significance level. Finally, we can remark the similar structure of correlations between each strategy and its Threshold Conjecture starting with the highest negative value that end up being positive as the Threshold Belief is further away from the decision node. In Section 4 we delve into the significant correlations between the strategies and the future Threshold Beliefs when we present avenues for future research.

$\bar{\mu}^1$	*First*	$\bar{\mu}^2$	*Second*	$\bar{\mu}^3$	*Third*	$\bar{\mu}^3$	*Pass*
$\bar{\mu}^1(1)$	**-0.4036*****						
$\bar{\mu}^1(2)$	0.0516**	$\bar{\mu}^2(2)$	**-0.1742*****				
$\bar{\mu}^1(3)$	0.2564***	$\bar{\mu}^2(3)$	-0.1116***	$\bar{\mu}^3(3)$	**-0.1491*****		
$\bar{\mu}^1(\infty)$	0.4140***	$\bar{\mu}^2(\infty)$	0.2066***	$\bar{\mu}^3(\infty)$	0.1491***	$\bar{\mu}^3(\infty)$	**-0.1491*****

Note: $^{*}p < 0.10$; $^{**}p < 0.05$; $^{***}p < 0.01$

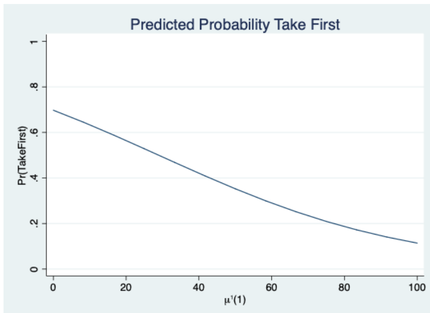
Table 2: Correlation Between (Updated) Threshold Beliefs and Behavior

To provide a more formal econometric analysis, we estimate logit models of each binary decision on the corresponding Threshold Belief, i.e. $\bar{\mu}^m(m)_{m=1,2,3}$. Table 3 in Appendix C presents the estimation results. Since the estimations corroborate the results in Table 2, we only include in the main text Figure 3 that visualizes the estimated probabilities of each strategy for changes in the contemporaneous Threshold Beliefs.

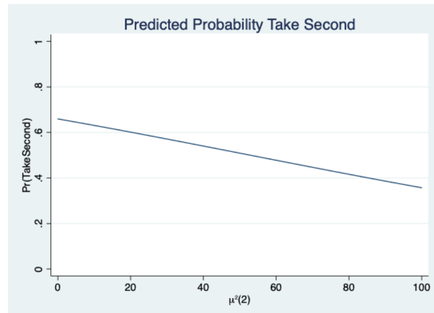
Figure 2 delivers the following observations. First, roughly 70% of the subjects would take at each decision node if the (Updated) Threshold Belief is 0, that is, if the Threshold Belief that player 2 would choose *Take* in the following decision node is 0. Following the interpretation shown, all subjects would present a subjective belief higher than the Threshold Belief, i.e. $\mu_i^1(1) \geq \bar{\mu}^1$ for all subjects, and then *Take* would be the unique best-response. Consequently, we can expect higher rates of *Take* in this case. If all subjects in the pool act following our proposed methodology, we could expect a predicted probability of *Take* equal to 100%.

Second, the sensitivity of the predicted probabilities to changes in Threshold Beliefs decreases as the game progresses. This is reflected in the slope of the predicted probability curve that becomes flatter as we move from Figure 3a to 3c. So, the subjects in the later decision nodes are less sensitive to changes in the Threshold Beliefs. In this line, mention that the average marginal effects of the explanatory variables in each model decreases by 50%, for the estimated logit models for *Take First* and *Take Third*, respectively. Note that, since the actions in the third node, *Take Third* and *Pass Always* are complementary, the estimated models, as well as the probabilities derived from them, are symmetric in Figure 3 and Table 2.

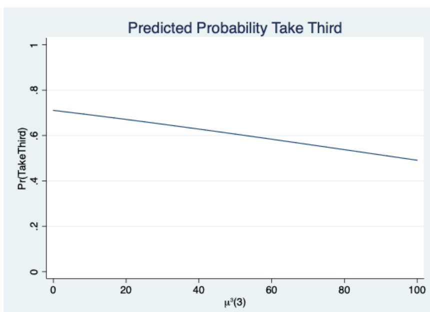
Last, the predicted probability to *Take* at each node in the hypothetical case of a (Updated) Threshold Belief equal to 100, in percentages, increase at later stages. As indicated by our methodology, if an (Updated) Threshold Belief would at some decision node be 100, nobody should play *Take* in this node. However, the estimated percentage in such a case is 15%, somehow higher than 0%, for the first node and increases up to 50% in the third node. This indicates that our methodology would explain worse the observed behavior in advanced nodes of the game.



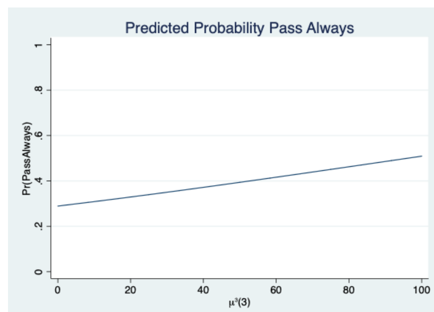
(a) Take First



(b) Take Second



(c) Take Third



(d) Pass Always

Figure 3: Predicted Probabilities of Strategies from Changes in its SizeBAP

All the results presented so far corroborate that the (Updated) Threshold Belief are significant predictors of the behavior.In the following, we quantify to what extent our methodology explains the observed behavior.As a first step, we evaluate to what extent our estimated logit models predict the observed individual choices. We define a correct estimation if, at the individual level, two conditions hold: (i) an estimated model predicts that a subject takes a particular action in a particular decision node with a probability higher than 50% and (ii) the observed decision is this particular action. The precision of the estimated models are 16.94%, 28.21% and 18.67% for *Take First, Take Second* and *Take Third*, respectively.

Previous literature (see Cooper, DeJong, Forsythe and Ross, 1996; Costa-Gomes, Crawford and Broseta, 2001; Garcia-Pola, Iriberri and Kovářík,2020; among many others) shows that individuals cannot be classified in an homogeneous decision making process. If that is so, we cannot expect that all subjects act following our belief-based approach. Therefore, we assess that Threshold Conjectures predict significantly the behavior of a subject if the prior Threshold Belief affects the observed behavior at less 10% significance for the estimated logit model in the first decision node. To do so, we estimate individually the logit model regressing the dummy for *Take First* on the prior Threshold Belief $\bar{\mu}^1(1)$ using the 16 decisions made by each individual. We carry out the classification procedure using the first decision node where there is a clearer relationship between the decision to enter the game or not and the prior Threshold Belief. We obtain that for 43, out of the total 71, subjects the effect is significant at 10% confidence level.[17] Precisely, for 26.76% of the subjects the *p-value* is equal or less than 1%, for 47.89% is equal or less than 5% and for 60.56% the significance level is less or equal than 10%.[18] Thereupon, our novel belief-based approach can explain significantly the differences across both subjects and CGs and, additionally, there are evidences that an important part of the decision makers may follow this methodology.

4 Conclusions and Future Research

In this study, we show that cooperation in CGs does not have to be puzzling. We theoretically present that cooperation in such games can be explained with a minimum suspicion about CBR. We further propose a novel general subjective belief-based approach that links p-rationalizability with the payoffs structure of CGs. We find that Threshold Conjectures, which induce indifference between possible strategies, serve as predictors of cooperation. Our methodology can explain the intra-subjects heterogeneity behavior, as a consequence of changes in the Threshold Beliefs of the CGs, and the between-subjects heterogeneity behavior, as a consequence of changes in the subjective beliefs. The reported results in this paper exhibit a well explanation of the lab data at both the population level and the individual level.

In the following, we outline three avenues for future research.

Design new CGs. In Section 3 we use CGs from Garcia-Pola, Iriberri and Kovářík (2020) which are designed for a different purpose purpose. There, these games were not designed to test our methodology. In order to test our theory further, we should design a set of alternative CGs, to exploit the ability of Threshold Conjectures to target causality, that systematically manipulate the magnitude and the evolution of the Threshold Conjectures.

Forward Looking. The framework presented in this paper explains the behavior in a

[17] There are 5 subjects who had never chose *Take First* so there is no estimated model for them.

[18] Table 10 in Appendix C displays both the cumulative and the cumulative relative frequencies for some ranks of *p-values*

particular node solely in function of the Threshold Belief in this node and disregards the Threshold Belief in later decision nodes. However, Table 2 and Table 5 in Appendix B show that all the future Threshold Beliefs, at every decision node, are strong predictors of the cooperation rates suggesting some king of *Forward Lookingness*. Future research should propose a theory linking the behavior in a particular node to payoffs in later stages of the game.

Strategy Uncertainty. Recent literature emphasizes in the ability of strategic uncertainty to organize the observed behavior in games that reflect the tension between payoff maximization and sequential rationality. Mostly of the papers applies the *Basin of Attraction* notion to the prisoners dilemma (Embrey, Fréchette and Yuksel, 2018; Dal Bó and Fréchette, 2011) but also to the CG (Healy, 2017). We can show that the first element of $\bar{\mu}^m$ at each decision node, i.e. $\bar{\mu}^m(m)_{m=1,2,3}$ for the case of CGs with $n = 3$, coincides with the *size of the Basin of Attraction to Pass*, which is the size of the set of beliefs such that cooperation is appealing. Therefore, their approach seems to be a subset of our more general methodology. As a result, further research should establish a more formal connection between our approach and this literature and between CGs and the games analysed in the above cited studies.

References

Aumann, Robert J. (1998). "On the centipede game". *Games and Economic Behavior*, **23**, 97–105.

Battigalli, Pierpaolo (1997). "On rationalizability on extensive-form games". *Journal of Economic Theory*, **74**, 40–61.

Battigalli, Pierpaolo and Marciano Siniscalchi (2002). "Strong belief and forward induction reasoning". *Journal of Economic Theory*, **106**, 356–391.

Battigalli, Pierpaolo and Marciano Siniscalchi (2003). "Rationalization and incomplete information". *The B.E. Journal of Theoretical Economics*, **3**, 1–46.

Bigoni, Maria, Marco Casari, Andrzej Skrzypacz and Giancarlo Spagnolo (2015). Time horizon and cooperation in continuous time. *Econometrica*, **83**(2), 587–616.

Bornstein, Gary, Tamar Kugler and Anthony Ziegelmeyer (2004). Individual and group decisions in the centipede game: Are groups more "rational" players? *Journal of Experimental Social Psychology*, **40**(5), 599–605.

Brandenburger, Adam, Alexander Danieli and Amanda Friedenberg (2019). Identification of reasoning about rationality.

Cooper, Russell, Douglas V DeJong, Robert Forsythe and Thomas W Ross (1996). Cooperation without reputation: Experimental evidence from prisoner's dilemma games. *Games and Economic Behavior*, **12**(2), 187–218.

Costa-Gomes, Miguel, Vincent P Crawford and Bruno Broseta (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica*, **69**(5), 1193–1235.

Dal Bó, Pedro and Guillaume R Fréchette (2011). The evolution of cooperation in infinitely repeated games: Experimental evidence. *American Economic Review*, **101**(1), 411–29.

Embrey, Matthew, Guillaume R Fréchette and Sevgi Yuksel (2018). Cooperation in the finitely repeated prisoner's dilemma. *The Quarterly Journal of Economics*, **133**(1), 509–551.

Fey, Mark, Richard D. McKelvey and Thomas R. Palfrey (1996). "An experimental study of constant-sum centipede games". *International Journal of Game Theory*, **25**, 269–287.

Friedman, Daniel and Ryan Oprea (2012). A continuous dilemma. *American Economic Review*, **102**(1), 337–63.

Garcia-Pola, Bernardo, Nagore Iriberri and Jaromír Kovářík (2020). Non-equilibrium play in centipede games. *Games and Economic Behavior*, **120**, 391–433.

Goeree, Jacob K and Charles A Holt (2001). Ten little treasures of game theory and ten intuitive contradictions. *American Economic Review*, **91**(5), 1402–1422.

Healy, Paul J (2017). Epistemic experiments: Utilities, beliefs, and irrational play. Tech. rep., mimeo, Ohio State University.

Ho, Teck-Hua and Xuanming Su (2013). A dynamic level-k model in sequential games. *Management Science*, **59**(2), 452–469.

Hu, Tai-Wei (2007). "On p-rationalizability and approximate common certainty of rationality". *Journal of Economic Theory*, **136**, 379–391.

Kawagoe, Toshiji and Hirokazu Takizawa (2012). Level-k analysis of experimental centipede games. *Journal of Economic Behavior & Organization*, **82**(2-3), 548–566.

McKelvey, Richard D. and Thomas R. Palfrey (1992). "An experimental study of the centipede game". *Econometrica*, **60**, 803–836.

McKelvey, Richard D. and Thomas R. Palfrey (1995). "Quantal response equilibria for normal form games". *Games and Economic Behavior*, **10**, 6–38.

McKelvey, Richard D and Thomas R Palfrey (1998). Quantal response equilibria for extensive form games. *Experimental economics*, **1**(1), 9–41.

Monderer, Dov and Dov Samet (1987). "Approximating common knowledge with common beliefs". *Games and Economic Behavior*, **1**, 170–190.

Nagel, Rosemarie and Fang Fang Tang (1998). Experimental results on the centipede game in normal form: an investigation on learning. *Journal of Mathematical psychology*, **42**(2-3), 356–384.

Pearce, David G. (1984). "Rationalizable strategic behavior and the problem of perfection". *Econometrica*, **52**, 1029–1050.

Perea, Andrés (2014). Belief in the opponents future rationality. *Games and Economic Behavior*, **83**, 231–254.

Rapoport, Amnon, William E Stein, James E Parco and Thomas E Nicholas (2003). Equilibrium play and adaptive learning in a three-person centipede game. *Games and Economic Behavior*, **43**(2), 239–265.

Rosenthal, Robert W. (1981). "Games of perfect information, predatory pricing and chain-store paradox". *Journal of Economic Theory*, **25**, 92–100.

A Computing the prior Threshold Conjecture

To obtain the prior Threshold Conjecture of P1, $\bar{\mu}^1$, for the McKelvey and Palfrey's (1992) *exponentially increasing sum* CG we proceed backwards. At each decision node we make indifferent P1 between *Take* and any future strategy. Note that whenever necessary beliefs are updated via bayesian rule.

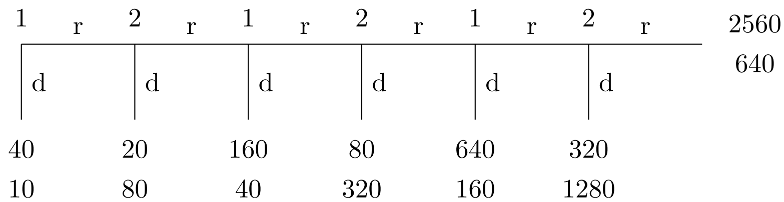


Figure 4: *Exponentially increasing sum Centipede Game*

$$\begin{cases} 640 = 320\frac{\bar{\mu}_1(3)}{\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} + 2560\frac{\bar{\mu}_1(\infty)}{\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} \\ 160 = 80\frac{\bar{\mu}_1(2)}{\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} + 320\frac{\bar{\mu}_1(3)}{\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} + 2560\frac{\bar{\mu}_1(\infty)}{\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} \\ 40 = 20\frac{\bar{\mu}_1(1)}{\bar{\mu}_1(1)+\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} + 80\frac{\bar{\mu}_1(2)}{\bar{\mu}_1(1)+\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} + \\ \qquad 320\frac{\bar{\mu}_1(3)}{\bar{\mu}_1(1)+\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} + 2560\frac{\bar{\mu}_1(\infty)}{\bar{\mu}_1(1)+\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty)} \\ \bar{\mu}_1(1)+\bar{\mu}_1(2)+\bar{\mu}_1(3)+\bar{\mu}_1(\infty) = 1 \end{cases}$$

To obtain the equations we proceed as follows:

- First equation. Payoff for *Take Third* equals to the expected payoff for any possible future final histories. That is, payoff for P1 when P2 chooses *Take Third* times the Threshold Belief $\bar{\mu}_1(3)$ over the probability to reach node 3 plus the payoff for P1 when P2 chooses *Pass Always* multiplied by the Threshold Belief $\bar{\mu}_1(\infty)$ over the probability to reach node 3.

- Second equation. Payoff for *Take Second* equals to the expected payoff for any possible future final histories. That is, payoff for P1 when P2 chooses *Take Second* times the Threshold Belief $\bar{\mu}_1(2)$ over the probability to reach node 2 plus the payoff for P1 when P2 chooses *Take Third* times the Threshold Belief $\bar{\mu}_1(3)$ over the probability to reach node 2, and so on.

- Third equation. Same procedure as the other two equations.

- Fourth equation. Restriction to get a full support conjecture.

B Estimated Models

	First	Second	Third	Pass
$\bar{\mu}(1)$	-0.0289***			
	(0.00216)			
$\bar{\mu}^2(2)$		-0.0125***		
		(0.00258)		
$\bar{\mu}^3(3)$			-0.00935***	0.00935***
			(0.00325)	(0.00325)
Constant	0.836***	0.661***	0.899***	-0.899***
	(0.120)	(0.139)	(0.200)	(0.200)
Observations	1,216	786	376	376

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Estimated Simple Logit Models for each Strategy

	First	*Second*	*Third*	*Pass*
$\bar{\mu}(1)$	-0.00960**	0.0147***	-0.00171	
	(0.00457)	(0.00469)	(0.00499)	
$\bar{\mu}^2(2)$	-0.0363***	5.37e-05	0.00257	
	(0.00685)	(0.00713)	(0.00744)	
$\bar{\mu}^3(3)$	-0.00394	-0.0296***	-0.0104*	
	(0.00524)	(0.00549)	(0.00577)	
Constant	3.039***	1.376***	0.929***	
	(0.226)	(0.237)	(0.246)	
Observations	1,216	1,216	1,216	1,216

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4: Estimated Simple Multinomial Logit Model

	First	First	First	Second	Second	Third
$\bar{\mu}(1)$	-0.0430***					
	(0.00323)					
$\bar{\mu}^2(2)$		-0.0474***		-0.0232***		
		(0.00415)		(0.00410)		
$\bar{\mu}^3(3)$			-0.0201***		-0.0248***	-0.0178***
			(0.00285)		(0.00346)	(0.00633)
Constant	1.936***	1.337***	0.409***	1.801***	1.755***	1.813***
	(0.187)	(0.154)	(0.121)	(0.250)	(0.207)	(0.463)
Observations	688	688	688	392	392	140

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Estimated Simple Logit Models, Forward Looking

C Additional Tables

	node 2			node 3	
CG	$\bar{\mu}^2(2)$	$\bar{\mu}^2(3)$	$\bar{\mu}^2(\infty)$	$\bar{\mu}^3(3)$	$\bar{\mu}^3(\infty)$
1	85.71	12.24	2.04	85.71	14.29
2	88.20	10.65	1.16	90.21	9.79
3	50.00	18.18	31.82	36.36	63.64
4	49.09	11.84	39.07	23.26	76.74
5	6.25	20.09	73.66	21.43	78.57
6	18.75	36.56	44.69	45.00	55.00
7	83.33	13.89	2.78	83.33	16.67
8	33.33	53.33	13.33	80.00	20.00
9	19.40	7.30	73.30	9.06	90.94
10	45.00	4.40	50.60	8.00	92.00
11	14.29	7.79	77.92	9.09	90.91
12	8.33	9.65	82.02	10.53	89.47
13	45.71	18.10	36.19	33.33	66.67
14	44.00	19.25	36.75	34.38	65.63
15	9.09	3.87	87.04	4.26	95.74
16	5.26	3.79	90.95	4.00	96.00

Table 6: Updated Threshold Conjectures from node 2 and 3 perspective

	First	Second	Third	Pass
$\bar{\mu}(1)$	-0.4036*	0.2317*	0.0903*	0.1471*
	(0.0000)	(0.0000)	(0.0016)	(0.0000)
$\bar{\mu}(2)$	0.0516	-0.0557	0.0246	-0.0241
	(0.0720)	(0.0520)	(0.3920)	(0.4012)
$\bar{\mu}(3)$	0.2564*	-0.1960*	-0.0356	-0.0489
	(0.0000)	(0.0000)	(0.2142)	(0.0885)
$\bar{\mu}(\infty)$	0.4140*	-0.2087*	-0.1157*	-0.1652*
	(0.0000)	(0.0000)	(0.0001)	(0.0000)

p-values in parentheses. $^*p < 0.01$

Table 7: Correlations between Threshold Beliefs and Strategies

	node 1				node 2			node 3
	$\bar{\mu}(1)$	$\bar{\mu}(2)$	$\bar{\mu}(3)$		$\bar{\mu}^2(2)$	$\bar{\mu}^2(3)$		$\bar{\mu}^3(3)$
$\bar{\mu}(1)$	1.0000			$\bar{\mu}(2)$	1.0000		$\bar{\mu}^3(3)$	1.000
	(-)				(-)			(-)
$\bar{\mu}(2)$	-0.4709	1.0000		$\bar{\mu}^2(3)$	-0.0026	1.0000	$\bar{\mu}^3(\infty)$	-1.0000
	(0.0000)	(-)			(0.9284)	(-)		(0.000)
$\bar{\mu}(3)$	-0.6829	0.4586	1.0000	$\bar{\mu}^2(\infty)$	-0.9109	-0.4103		
	(0.0000)	(0.0000)	(-)		(0.0000)	(0.0000)		
$\bar{\mu}(\infty)$	-0.8652	0.0320	0.3354					
	(0.0000)	(0.2642)	(0.0000)					

Note: *p-value* between brackets

Table 8: Correlation between (Updated) Threshold Beliefs

	$\bar{\mu}(1)$	$\bar{\mu}^2(2)$	$\bar{\mu}^3(3)$
$\bar{\mu}(1)$	1.0000		
	(-)		
$\bar{\mu}^2(2)$	0.6610	1.0000	
	(0.0000)	(-)	
$\bar{\mu}^3(3)$	0.4647	0.7956	1.0000
	(0.0000	(0.0000	(-)

Note: *p-value* between brackets

Table 9: Correlation between First Element of (Updated) Threshold Conjectures

p-value	Cum. Freq	Cum. Rel. Freq
≤ 0.005	13	18.31
≤ 0.010	19	26.76
≤ 0.050	34	47.89
≤ 0.100	43	60.56
≤ 0.250	57	80.28
≤ 0.500	59	83.10
≤ 0.750	68	95.77
≤ 1.000	71	100

Table 10: Frequencies of *p-value* used to Classify Individuals

D Centipede Games

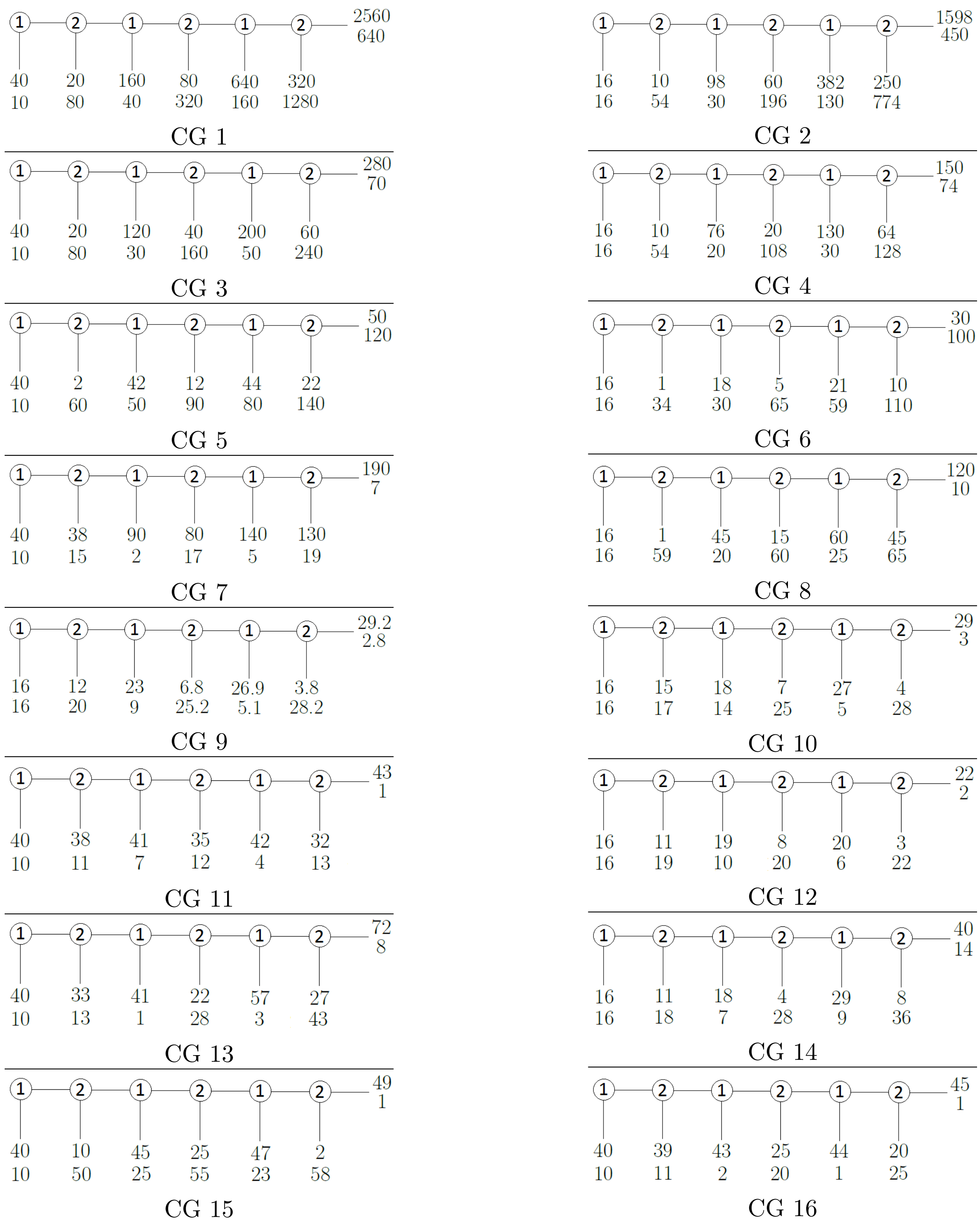


Figure 5: The 16 CGs Used in the Experiment.

E Additional Plots

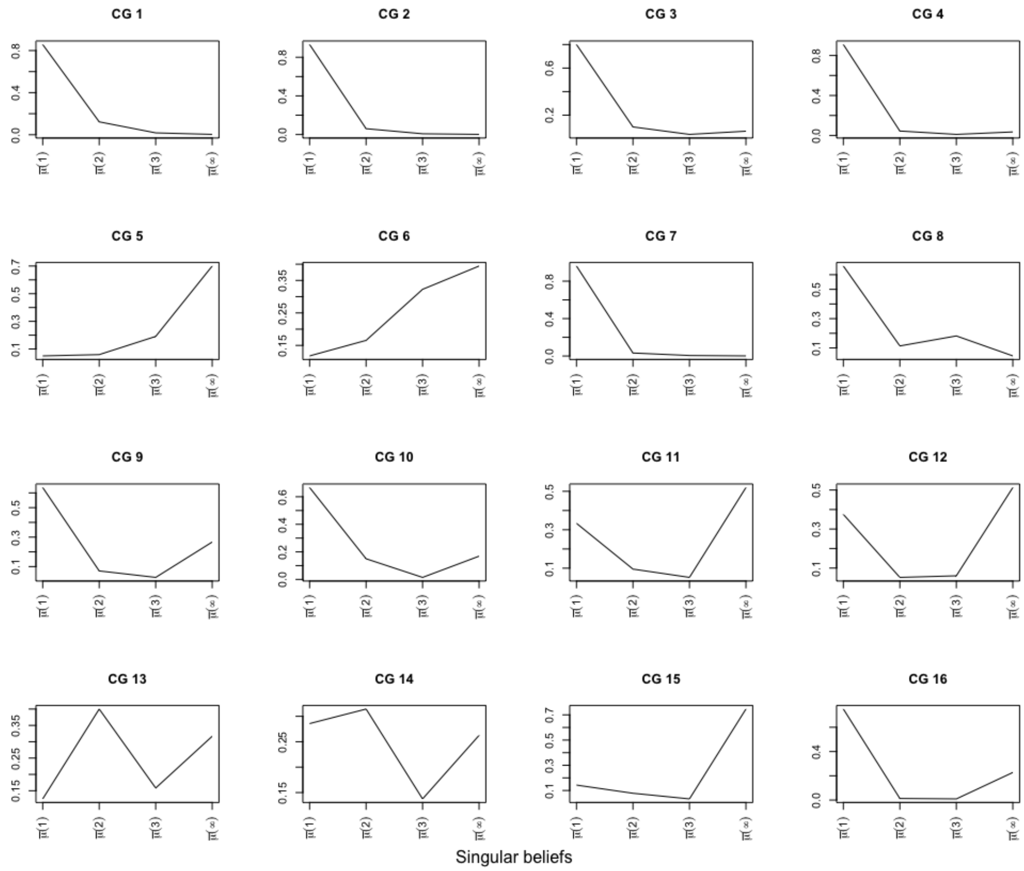


Figure 6: Evolution of prior Threshold Conjectures across CGs

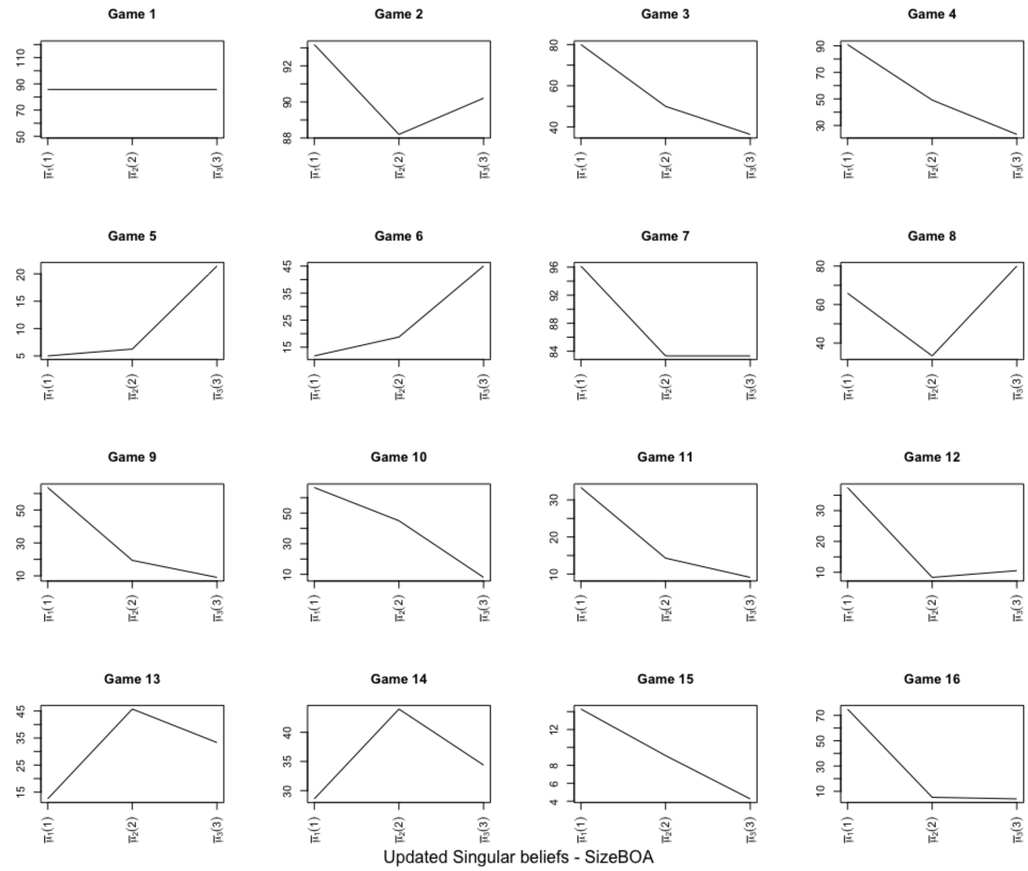


Figure 7: Evolution of Updated Threshold Beliefs across CGs

The Updated Threshold Beliefs plotted in this figure are the Threshold Beliefs that player 2 would play $s_2(m)_{m=1,2,3}$, from each node perspective m. That is, $\bar{\mu}_1^m(m)_{m=1,2,3}$, which correspond to second column of Table 1a and second and fifth columns of Table 6.

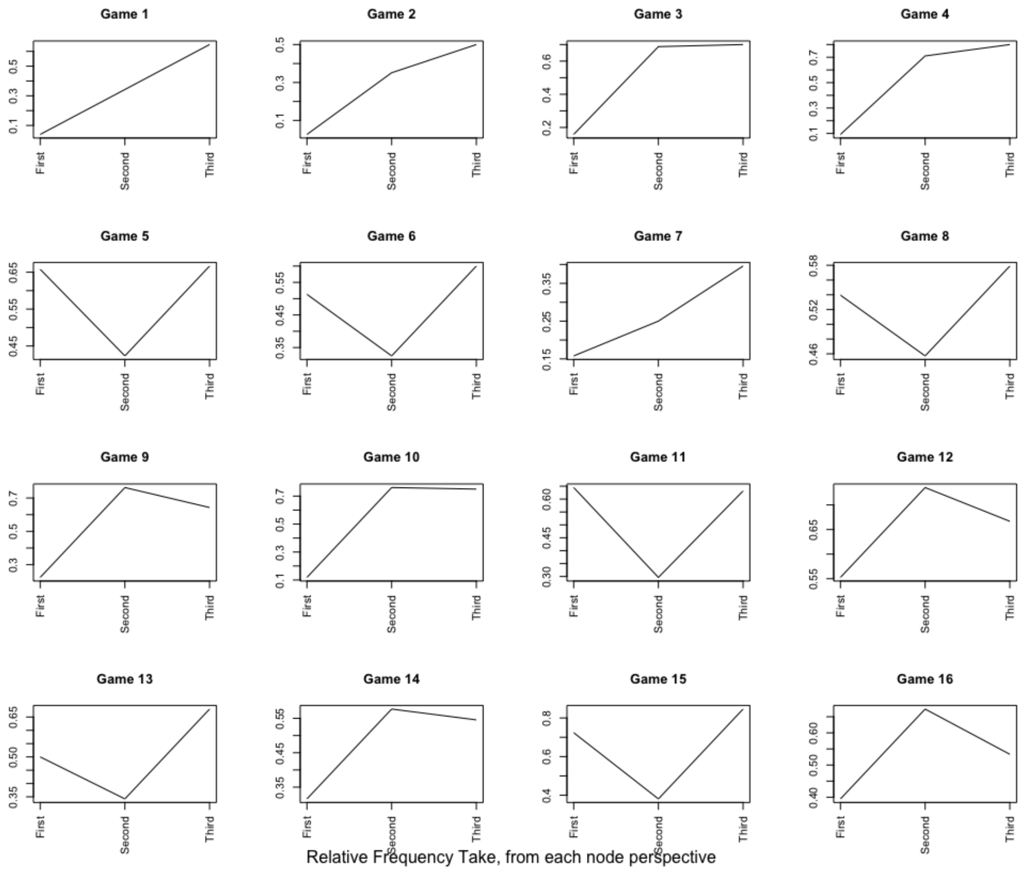


Figure 8: Evolution of Strategy's Frequencies across CGs

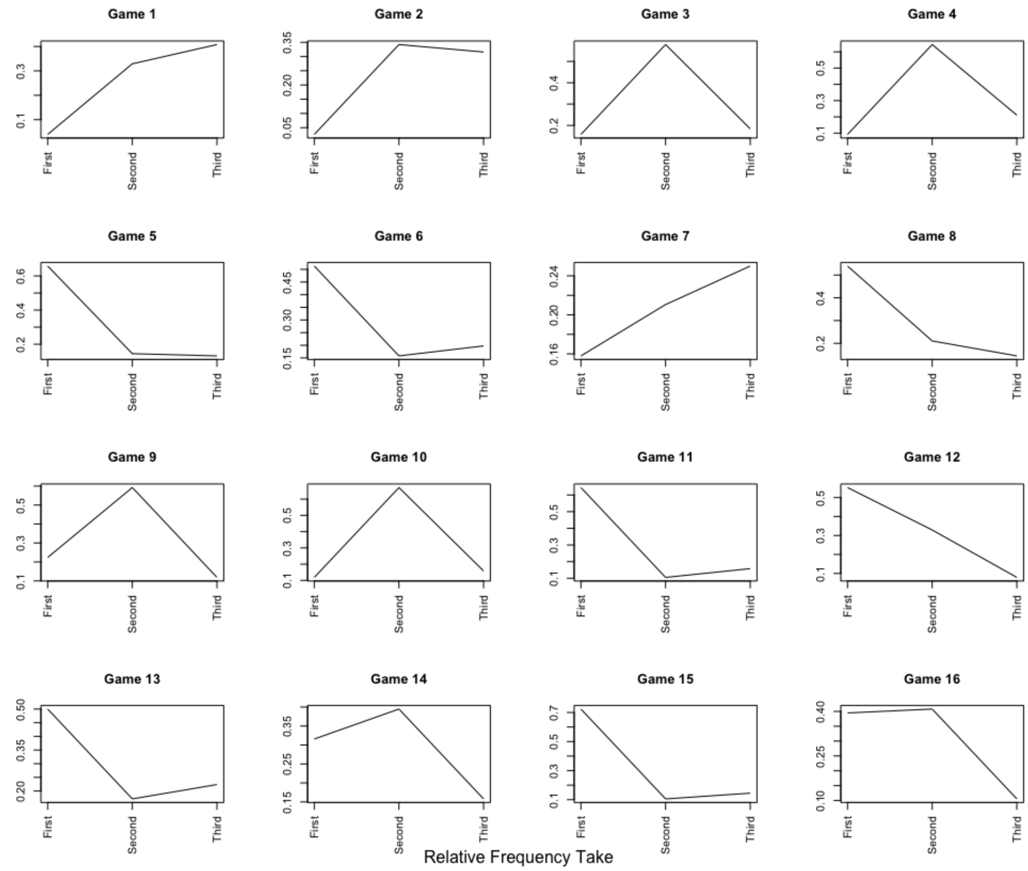


Figure 9: Evolution of Conditioned Strategy's Frequencies across CGs