Selten's Horse:
An experiment on sequential rationality*

Asha Sadanand† Nikolas Tsakas‡

April 11, 2023

Abstract

In a seminal paper, Selten (Selten, 1975) developed the game Selten's Horse to illustrate some aspects of rationality. In our experimental study of Selten's Horse we find that behaviour is not according to conventional sequential rationality based refinements. Some behaviour is better explained by Ideal Reactive Equilibrium (Sadanand, 2019), an extension of Manipulated Nash Equilibrium (Amershi, Sadanand, and Sadanand, 1985a,b,c). We provide a simple model of Level–k thinking incorporating the idea of virtual observability, where the players behave as if their subsequent opponents could in effect observe their actions. We find that participants' behaviour is somewhat explained by virtual observability, but this is limited to immediate successors and is more prevalent in the more complex version of the game. Moreover, at times we see first movers naïvely move towards their most-preferred equilibrium and follower players anticipate and exploit it by systematically reaching off-equilibrium outcomes that are favourable to them.

Keywords: sequential rationality; Selten's Horse, level-k, virtual observability.

JEL classification: C72, C73, C91, C92

Declaration of interests: none.

*We are grateful for insightful comments from colleagues at the University of Cyprus, University of Guelph, 30th International Conference on Game Theory, Wilfred Laurier University, 6th World Congress of the Game Theory Society, Curtin University, and University of Auckland. This research was funded by University of Cyprus Starting Grant 8037E-32806. Financial support is gratefully acknowledged.

†Department of Economics and Finance, University of Guelph, 50 Stone Road East, Guelph, ON N1G 2W1, Phone (519)824-4120 x58947. Fax +1(519)763-8497e-mail: asadanan@uoguelph.ca, Corresponding author.

‡Department of Economics, University of Cyprus, PO Box 20537, 1678 Nicosia, Cyprus, e-mail: tsakas.nikolaos@ucy.ac.cy.

1. Introduction

One of the most challenging frontiers in game theory, has been the prediction of behaviour when there are multiple Nash equilibria. In the absence of any method of ranking the merits of each equilibrium, it would baffle even the most rational players to somehow coordinate on one equilibrium. Non-equilibrium outcomes are very likely even if all players are using equilibrium strategies. This has lead to a burgeoning literature of esoteric refinements that require high levels of mathematical sophistication on the part of the players. Of course, the argument made is that the players choose *as if* they could work out all the mathematics, but there comes a point when even the *as if* could perhaps be far too demanding.

The importance of required sophistication on equilibrium selection is one main aspect of interest in this paper. Namely, we want to test whether, in the presence of multiple equilibria, experimental subjects will coordinate on equilibria predicted by several conventional refinements, which require sequential rationality; or they will be attracted by other more intuitive equilibria that seem more natural, but which are rejected by the conventional refinements.

Second is the impact of the order of play in equilibrium behavior. There is a large literature that focuses on the impact of presentation effects and of the order in which players move, both in normal form and in extensive form games. It is well documented (for instance, see Camerer, Knez, and Weber, 1996) that in several different environments, the timing of the moves might affect the outcome of a sequential game when earlier moves are not observable to later players. This is often attributed to two potential reasons: (i) First–mover advantage, which suggests that a player who plays first will choose the path of her most preferred equilibrium and this will in turn be followed by later players, (ii) Virtual observability, which suggests that players will anticipate the earlier players' moves and act as if they can observe their choices. This latter effect is generalized in a intutitively based refinement by Amershi, Sadanand, and Sadanand (1985a,b,c), Manipulated Nash Equilibrium, that was further developed and termed Ideal Reactive Equilibrium Sadanand (2019).

In addition to these equilibrium predictions, Guth, Huck, and Rapoport (1998) have considered the possibility that players who act first may steer the game to non–equilibrium outcomes that might be more favorable for them. In the Ideal Reactive Equilibrium theory, this incentive is not exclusive to first movers, but that all players will want to steer the continuation of play from whichever point it is their turn. However, moving first induces commitment, that can be subsequently exploited by later players. In a version of the game that we consider here, the existence of a payoff dominant equilibrium might induce one of the players to commit to a strategy that leads to this equilibrium, but can then be anticipated and exploited by subsequent players.

In our experiment, we focus on Selten's Horse (Selten, 1975),[1] a game that was designed by Selten to illustrate the Trembling Hand Perfect refinement. The game is explained in detail in Section 2. As it turns out, it has exactly the features that we need in order to test our theoretical hypotheses. There are two Nash equilibrium outcomes in pure strategies (with two connected components of Nash equilibria), and of these two equilibrium outcomes, one is rejected by all conventional refinements using sequential rationality, thus also rejected by Selten (1975). However, the equilibrium that is rejected, is exactly the one predicted by the Ideal Reactive Equilibrium (IRE), which in turn rejects the equilibrium predicted by all refinements requiring sequential rationality.

The reason for the difference between the predictions is that all of the conventional theories require rationality at all decision points as a means to remove incredible threats. However, this amount of rationality can be an overkill. It imposes optimal behaviour at some information sets where it is actually unimportant what the player plans to do, as none of their actions pose an incredible threat. Instead of applying such excessive rationality, IRE does not presume any specific behaviour at any unreached information sets, but instead requires all players to follow a process of elimination called *thought process dynamics*, whereby each player goes through all the possible *what ifs* (which are restricted in a specific way by the order of play) before they arrive at a decision of what to choose.

The experimental results provide some evidence of our main intuition. In the orginal order of play, we find that subjects not only do they not coordinate on the equilibrium that is predicted by conventional refinements, but they strongly avoid it. In fact, the most frequently observed outcome is the IRE predicted outcome. In our second treatment where we alter the order of play, IRE and conventional theory predictions coincide. However, here we do not see much of the play predicted by both theories.

In an attempt to better understand the observed behaviour, we construct an alternative model that keeps the main feature of IRE, that of virtual observability. Namely, we consider that players believe that there is a positive probability that their choices will be observed by subsequent players when they are not supposed to. In a sense, this describes a notion of probabilistic virtual observability. Furthermore, we consider a framework of hierarchical level-K thinking (see Stahl and Wilson, 1994)) to capture different levels of subject sophistication.

When testing the predictions of the alternative models, we find that subjects' behavior is indeed explained to some extent by virtual observability, but this is limited to immediate successors. That is, subjects believe that with some probability their immediate successor will be able to anticipate their move and best–respond to it. Notably, we find this behaviour to be more prevalent in the variant of the game in which the order is reversed, in which the estimated level of players' sophistication is lower. This does not reflect differences

[1] We accidentally departed from Selten's original game. As we shall see in one of the end nodes we have Player # 2 receive a payoff of 3, while in Selten's original game it was 2.

in the cognitive abilities of the subjects who participated in this treatment, but rather that the perceived complexity of that game is higher.

Furthermore, virtual observability is often believed to provide a first–mover advantage, as players who play first can lead their game towards their most preferred outcome. However, in the current framework, when the order of play is reversed, players who play last could anticipate this tendency of the first mover and exploit this commitment to their advantage. We find evidence that supports this claim, yet it is interesting that this tendency leads them to inferior results than what coordination to the payoff–dominant equilibrium would yield.

The results are to a large extent aligned with those of the recent experimental work on Selten's Horse by Neugebauer, Sadrieh, and Selten (2022). Although, their experimental design was very different, interestingly, the experimental observations are remarkably similar between the two papers. In fact, Neugebauer, Sadrieh, and Selten (2022) observe an even greater frequency of the IRE predicted behaviour, which combined with our result make it a reasonable candidate to explain behaviour in this type of games.

2. Literature Review

The tension between conventional sequential rationality predictions and IRE predictions for Selten's Horse is an instance of the more general observation that order of play makes a difference even in games where moves of earlier players many not be observed.

The first experimental results that suggested an order of play effect was the work by Cooper, DeJong, Forsythe, and Ross (1990, 1993, 1994). There, the authors investigated the battle of the sexes game to see how experimental subjects dealt with outside options. They presented subjects with sequential unobserved battle of the sexes games without outside options, with outside options that were credible, and with outside options that were not credible. They expected to find that only in the games where the outside option was credible would you find a large propensity toward the Player 1 preferred equilibrium. They were puzzled to find that it did not matter whether the outside option was credible, or whether there even was an outside option. Players seemed to predominantly play the Player 1 preferred Nash equilibrium. Camerer, Knez, and Weber (1996) set out to test whether real subjects played according to the Amershi, Sadanand and Sadanand Mapnash equilibrium for weak link games such as the stag hunt. They found that in later rounds, their results confirmed a first mover advantage, which they called "virtual observability". Muller and Sadanand (2003) have found that subjects often choose in accordance with "virtual observability," particularly for simple games. Brandts and Holt (1992) and Rapoport (1997) have also found order of play effects. In addition, Schotter, Weigelt, and Wilson (1994) gave important results on how experimental subjects perceive order of play based on the presentation of the games.

The question that this growing strand of the experimental literature poses is: Exactly how does order of play affect outcomes? Indeed, Hammond (2008) stated that the conventional assumption of normal form invariance may be “unduly restrictive” and emphasized the need for a systematic theory of how order of play affects outcomes. Huck and Müller (2005) conducted some experiments and found that the opportunity to burn money confers a substantial advantage to first movers, and they too urge that “we make theoretical advances to understand the role of physical timing and first-mover advantages in games.” One such theory is IRE, and in this paper we explore in the context of Selten’s Horse whether experimental subjects adhere to conventional sequential rationality, or are influenced by order of play.

In addition to the order of play literature, another important piece that has already been mentioned is the experimental work relating to Selten’s Horse by Neugebauer, Sadrieh, and Selten (2022).

We recognize that experiments may not be able to provide a conclusive result as to whether IRE is played or the conventional theory prediction. There has been some literature addressing the general question of whether experimental participants conform to Nash equilibrium refinements. Gale, Binmore and Samuelson (1995) find that evolutionary models seem to describe behaviour well. Some of the earliest expeiments on ultimatum games were conducted by Guth, Schmittberger, and Schwarze (1982) and Guth and Kocher (2013) provides an indepth discusssion of the results from experimental research on the ultimatum game. Charness and Defwenberg (2006) study experimental subjects in a principal and agent relationship, and utilize the device of communication to create an obligation to comittment, that steers players to (non-equilibirum) committment outcomes. The main conclusion of this literature is that except in very simple games, experimental subjects generally do not exhibit equilibrium behaviour. What remains then, is to tease out whether this non-conformance to equilibirum behaviour is due to the subjects’ inability to compute or identify or reason out a particular refinement equilibirum, or whether it is that the subjects never did seek to align their actions with this refinement but instead their behaviour is motivated by some other process.

In Section 3 we present our theoretical analysis on Selten’s Horse and its variant. As neither game is very simple, it is understood that we may not see much equilibrium behaviour. We then construct an alternative model that allows for some virtual observability. We explain the structural model that will be used for estimation. In Section 4 we present experimental design and the testable hypothesis. In Section 5, we present the experimental results, in Section 6 we discuss our results with respect to some alternative behavioral rules and Section 7 concludes the paper.

3. Theory

3.1. Preliminaries: Selten's Horse and its variant

We develop our analysis around the game commonly known as Selten's Horse. Selten (1975) developed this game of imperfect information to illustrate the inadequacy of subgame perfection when there are non-trivial information sets. In this context, he introduced extensive form trembling hand perfect equilibrium (as distinct from normal form trembling hand perfect) as an extension of his subgame perfect concept (Selten, 1965). The game with the players' payoffs[2], is depicted in Figure 1. It is a three–player game (players denoted as P1, P2 and P3), in which P1 and P2 have fully aligned preferences.

We consider also an alternative version of this game in which the order of some moves is reversed, which is presented in Figure 2 and we will refer to it as variant of the game throughout the paper. Note that, the move reversal preserves the normal form; $P1$ and $P2$ are still unaware of $P3$'s action, despite this having taken place already. It is apparent that the Nash equilibria of the two games will be the same since the normal form equivalent of both games is, given in Figure 1 is the same.

Let $p, r, q \in [0, 1]$ denote the probabilities with which $P1$ plays D, $P2$ plays d and $P3$ plays L, respectively. The game has two pure strategy Nash equilibria are DcL giving payoffs of (3, 3, 2) and CcR giving payoffs of (1,1,1). There are also two sets of Nash equilibria that involve mixed strategies by some player: $(D, r \leq 2/3, L)$ and $(C, c, q \leq 1/4)$. It is important to notice here that both pure strategy Nash equilibria prescribe action c to $P2$ and even in all mixed equilibria action c is expected to be played quite often. This game is particularly challenging for Player #3 because he is not only uninformed about which history transpired before his move, but the possible histories involve different sequences of players moving.

3.2. Equilibrium Selection

All refinements of that utilize sequential rationality,[3] select CcR as their prediction in both games. The reasoning is that with DcL, according to the equilibrium $P2$ is required to choose c with a high enough probability, although he never actually plays. But this is not an optimal choice, since if $P2$ ever did get to play, knowing that in equilibrium $P3$ is moving L, it would be optimal for $P2$ to choose d instead of c. This is a violation of sequential rationality. On the other hand, CcR, accords with sequential rationality, since $P2$'s choice of c is completely rational, as now $P3$ plans to play R. The rest of the players are also

[2] As noted earlier, there is a slight discrepancy in the payoffs. While the orginal Selten's Horse had the payoffs 3,2,2 for the path D(·)L, in the version that we used for our experiment we accidentally used the payoffs 3,3,2. As it happens this slight error does not in any way change the predictions of each of the theories, or really any aspect of the game.

[3] Includes Selten's extensive form Trembling Hand Perfect (Selten, 1975), Perfect Bayesian Equilibrium (Fudenberg and Tirole, 1991), Weak sequential equilibrium, Sequential equilibrium (Kreps and Wilson, 1982), and all forward induction equilibrium concepts such as those used in signaling games.

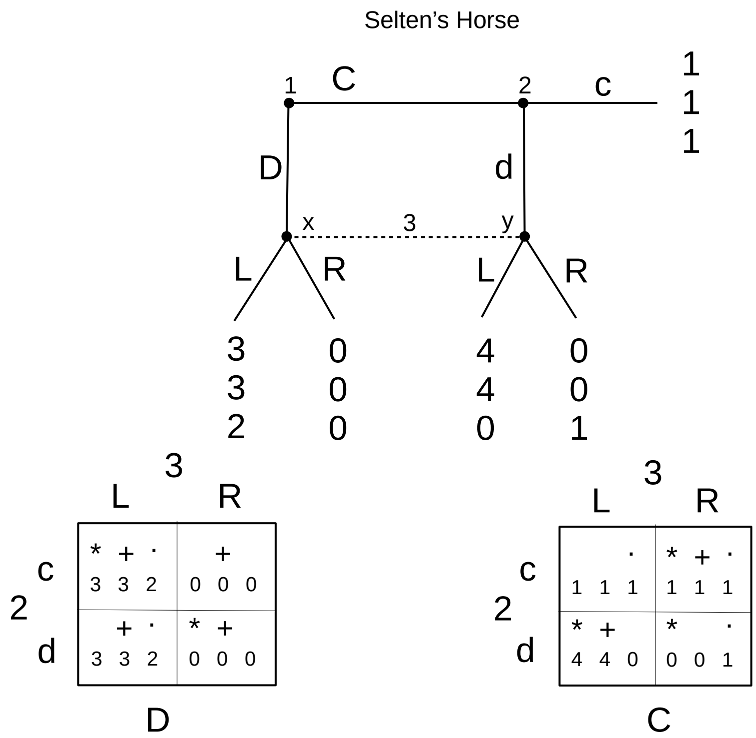


Figure 1: Selten's Horse

making rational choices based on their beliefs and the remaining players' equilibrium actions, making CcR sequentially rational. Note that this reasoning is not affected by the order of play.

Now, in contrast, we consider also an alternative equilibrium refinement, Ideal Reactive Equilibrium Sadanand (2019), which incorporates the order of play. Ideal Reactive Equilibrium (IRE) is based on the idea of "virtual observability": each player goes through a process of imagining that any arbitrary number of subsequent players are able to observe his even if in the original game they could not. In some special games it means that a Nash equilibrium survives the IRE refinement if it is the subgame perfect equilibrium outcome of the related game where all information sets are removed.[4] Selten's Horse is such a game. However, in general, Ideal Reactive Equilibrium also considers processes in which only some information sets are removed and not others. It is apparent that for IRE the order of play is important. The thought process dynamics that is utilized gives importance to the order of play, with earlier players having potentially greater control over the game, by making moves first that eliminate large parts of the game. However, there is a commitment aspect whereby after a player has moved he relinquishes control of the game, and the rest of what unfolds is completely in the hands of the remaining players.

In the original Selten's Horse game IRE predicts the DcL outcome, contrary to all refinements that

[4]For a more detailed presentation of IRE see Sadanand (2019).

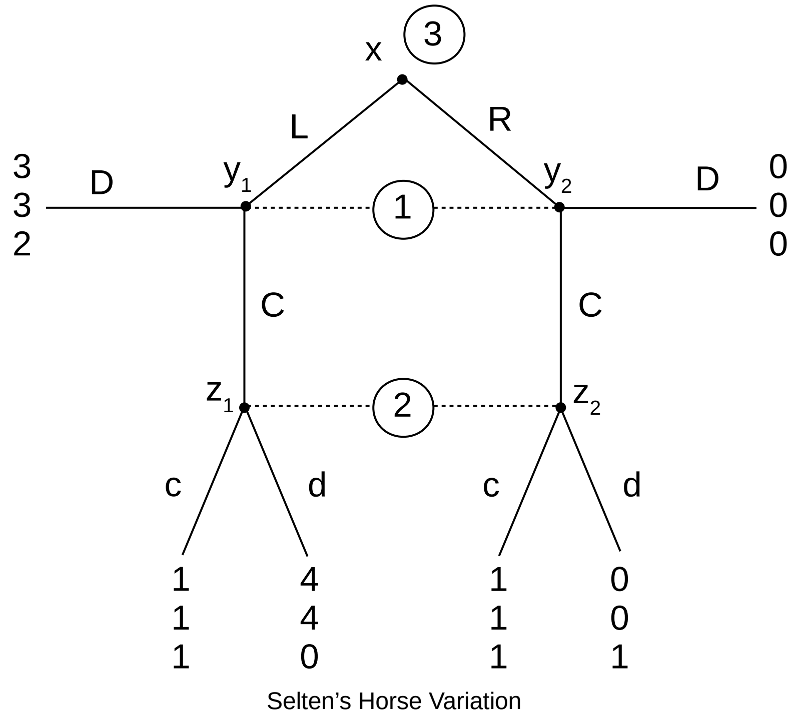


Figure 2: Selten's Horse with the order of play varied

require sequential rationality, whereas in the other variant of Selten's Horse IRE predicts the CcR outcome, which is in agreement with sequential rationality.

For the original game, the reasoning is as follows: Looking at the terminal nodes after $P2$'s move, the only one that could have been of interest to $P1$ is the CdL outcome giving payoffs of (4, 4, 0). But all players realize that one is not attainable in equilibrium, because if $P3$ thought that the others were playing D and d he would not play L. The rest of the terminal nodes after $P2$'s move have very low payoffs for $P1$, and so Ideal Reactive Equilibrium predicts that $P1$ would choose D and eliminate half of the game including $P2$'s move. All the players can anticipate this choice from knowing the game tree and by putting themselves in $P1$'s shoes. $P3$ knowing (or correctly anticipating) $P1$'s choice, would respond with L, ending the game and giving everyone the payoffs (3, 3, 2).

For the variant, if $P3$ moves L he succeeds in eliminating a large part of the game, but also gives up control of the game, because once a move is taken, he cannot take it back. If the remaining players correctly anticipate that $P3$ has moved L then it is almost as if they play a subgame from y_1 onward. Clearly the equilibrium of that continuation play is Cd, which gives a low payoff for $P3$. IRE posits that $P3$ will anticipate this, and not move L in the first place.

With games of perfect information, subgame perfect equilibrium (Selten, 1965) gives us mostly unique equilibria by eliminating equilibria that are supported by incredible threats. If we presume that the main purpose of conventional refinements for dynamic incomplete information games is to extend the idea of eliminating incredible threats in these games, then we should look to see if there were any incredible threats. Looking at the CcR component, $P3$ is not making an incredible threat with R, since there are beliefs over his information set that make R optimal. So this outcome is certainly fine from the perspective of incredible threats. Looking at the DcL component, $P2$'s choice of c is not really an incredible threat, since all players know that the CdL outcome is not an equilibrium, and there is no equilibrium where $P2$ plays d. Thus, this "rationality" that is required of $P2$, that motivates conventional theory to reject DcL, seems excessive.

One thing to note about the variant is if there is only partial IRE thinking whereby $P3$ continues to play L (possibly because of level-1 thinking that $P1$ and $P2$ are level-0 and just randomize equally), and if $P1$ and $P2$ realize this, it will induce them to take control of the game after L and to reach the terminal node CdL which gives them their maximum payoff, and a payoff of 0 to $P3$. This is a behavioural feature that would require incomplete IRE thought process thinking on the part of $P3$. We shall see a general discussion of partial IRE thinking in the next section.

3.3. An alternative model of (partial adherence to) virtual observability

Testing for equilibrium selection in the lab is often tough, as it presumes that play will mostly converge to some equilibrium outcome and one would be able to compare frequencies of occurrences of different equilibria. However, the complexity of a game or other behavioural factors might preclude players' choices to coincide with the predictions of equilibria. We attempt to tackle this problem in two ways. First, we construct a simple model of *probabilistic virtual observability* according to which players maximize their expected utility considering that some (or all) of the subsequent players will be able to observe their choice, even if they are not supposed to according to the game structure. Second, we incorporate some form of bounded rationality, which is discussed in the next subsection

More specifically, we consider that each player maximizes his expected utility from each of the two actions, believing that:

- the player who follows in the order of play will be able to observe his action when not supposed to with probability $\alpha \in [0, 1)$.
- both players who follow in the order of play will be able to observe his action when not supposed to with probability $\beta \in [0, 1)$.
- no information set will be removed compared to the game's structure with probability $1 - \alpha - \beta$.

The model captures the basic feature of IRE, which is that players consider that others might virtually observe (or predict) their actions, even though they are supposed to. Observe that $\alpha = \beta = 0$ brings us back to the original game and as β approaches 1, the game approaches that of perfect information, with the limitation that each player still makes a single choice and cannot condition this on the node he is at. It is also apparent that in this framework the order of moves will be important for determining the expected utility of each action. The usefulness of this model is that it allows us to quantify the extent to which the players are affected by virtual observability, in a way that accounts also for limited levels of sophistication in the thought process required by IRE.

Namely, for the original game the expected utilities are the following:

- $EU_1(D) = 3q(1-\alpha-\beta) + 3(\alpha+\beta)$ and $EU_1(C) = 4rq(1-\beta) + (1-r)$,
- $EU_2(d) = 4q(1-\alpha-\beta)$ and $EU_2(c) = 1$,
- $EU_3(L) = 2p + (1-p)(1-r)$ and $EU_3(R) = 1-p$.

whereas for the variant the expected utilities are the following:

- $EU_1(D) = 3q$ and $EU_1(C) = (1-\beta)(4rq + 1 - r) + \beta(3q+1)$,
- $EU_2(d) = 4q$ and $EU_2(c) = 1$,
- $EU_3(L) = (1-\alpha-\beta)\left[2p + (1-p)(1-r)\right] + \alpha\left[2p_l + (1-p_l)(1-r)\right]$ and
  $EU_3(R) = (1-\alpha-\beta)(1-p) + \alpha + \beta$.

where p_l is the probability with which P3 believes that P1 will play C upon virtually observing L, when P2 will not be observing P1's action. Given that P3 assumes that P1 will be maximizing his own expected utility, we get that $p_l = 1$ if $r < 2/3$, $p_l = 0$ if $r > 2/3$ and $p_l \in [0,1]$ if $r = 2/3$. If P1 virtually observes R he plays always C.[5]

The equilibria of the two games will not coincide anymore and will, of course, depend on α and β. Interestingly, there are equilibria for even small (but non–negligible) values of α and β such that playing d can be an equilibrium strategy for P2. The complete set of equilibria for both games is presented in the Appendix.

3.4. Bounded Rationality

The second tool we use in order to better understand the features of players' behavior in such a complex environment is to introduce a boundedly rational framework of belief formation, which will be more likely to describe accurately the behavior of experimental subjects. In particular, we consider a framework of

[5] D is weakly dominated by C and in fact gives a strictly lower payoff for any $r < 1$.

hierarchical Level–k thinking (Nagel, 1995; Stahl and Wilson, 1994) according to which players of Level–0 randomize uniformly across actions and, for all levels $k \geq 1$, Level–k players consider that the society consists of players of levels at most $k-1$. Interestingly, even the simplest deterministic model at which Level–k players consider a population consisting entirely of Level–$(k-1)$ predicts quite different behavior from the prescription of Nash equilibria (see Table 1). Namely, it is striking that P2 of up to Level–3 is expected to play d, which is in stark contrast with equilibrium predictions. This feature is still present if we consider the same process in the alternative model for both variants of the game.

This brings into light another feature of the games which is completely disregarded by both equilibria. This is the fact that observing the outcome CdL, which yields the maximum payoff to P1 and P2 and the minimum payoff to P3, might not actually be unexpected. It is enough for P3 to believe that there is a sufficiently high probability that P1 will play D. Inducing virtual observability with some probability has adverse effects in the two versions of the game. In the original version, P1 and P2 are the ones that become more careful when they believe that P3 might observe their action, thus leading P1 to be more inclined towards D. On the contrary, in the variant, P3 is the one who needs to be careful, as the possibility of him being observed to have chosen L might lead him towards a bad outcome. Yet, these arguments are under the assumption that the level of sophistication in players' choices is the same in both versions, which might not necessarily be true.

	$\alpha=\beta=0$			$\alpha,\beta\geq 0$, Original			$\alpha,\beta\geq 0$, Variant		
Level	p	r	q	p	r	q	p	r	q
Level–0	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2	1/2
Level–1	1/2	1	1	1	1	1	0	1	1
Level–2	0	1	1	0	1	1	0	1	0
Level–3	1	1	0	0	1	0	0	0	0
Level–4	1/2	0	0	1	0	0	0	0	1
Level–5	0	0	1	0	0	1	1	1	1

Table 1: Behavior according to Level–k thinking, when players of Level–0 mix uniformly and players of Level–k best respond to players of Level–$(k-1)$. First column correspond to the original model without virtual observability. Second and third column correspond to the alternative model of the original game and its variant respectively. For the two last columns we assume $0 < \alpha+\beta < 1/3$, $\alpha > \beta$ and $\beta < 1/4$.

In our analysis, we use the slightly more involved framework of Stahl and Wilson (1994). More specifically, we consider a population that consists of three distinct types of players (Level–0, Level–1 and Level–2), where Level–0 players mix uniformly, Level–1 players consider that the population consists solely of Level–0 players and Level–2 players consider that the population consists both of Level–0 and Level–1 players.

In addition to this, players' choices are subject to errors, which occur proportionally to the relative utility of the two actions of each player. Therefore, if $EU_i(j|p, r, q, \alpha, \beta)$ is the expected utility that player i gets from choosing action j, we consider that a player P_i of Level–1 chooses action j in round t with probability:

$$p_{it}^1(j) = \frac{exp\left(\gamma_1 EU_i(j|p^0, r^0, q^0, \alpha, \beta)\right)}{\sum_l exp\left(\gamma_1 EU_i(l|p^0, r^0, q^0, \alpha, \beta)\right)}$$

where $p^0 = r^0 = q^0 = 1/2$ are the probabilities with which each Level–0 player chooses action D, d and L respectively and γ_1 describes the precision of the player's calculations. Let p^1, r^1, q^1 be the probability with which a player of Level–1 chooses D, d and L respectively.

Given this, players of Level–2 form beliefs about the expected choices of others, considering that all other player are either of Level–0 or of Level–1. These beliefs can be approximated by a similar expression, with a potentially different level of precision μ_1.

$$\hat{p}_{it}^1(j) = \frac{exp\left(\mu_1 EU_i(j|p^0, r^0, q^0 \alpha, \beta)\right)}{\sum_l exp\left(\mu_1 EU_i(l|p^0, r^0, q^0, \alpha, \beta)\right)}$$

Let $\hat{p}^1, \hat{r}^1, \hat{q}^1$ be the belief of a player of Level–2 regarding the probability with which a player of Level–1 chooses D, d and L respectively. These beliefs can now be used to define the probabilities with which players of Level–2 choose each action.

$$p_{it}^2(j) = \frac{exp\left(\gamma_2 EU_i(j|\hat{p}^1, \hat{r}^1, \hat{q}^1, \alpha, \beta\right)}{\sum_l exp\left(\gamma_2 EU_i(l|\hat{p}^1, \hat{r}^1, \hat{q}^1, \alpha, \beta\right)}$$

and as before p^2, r^2, q^2 is the probability with which a player of Level–2 chooses D, d and L respectively.

Using these values, we could now turn to the data to estimate both the proportions of players of different types/levels within the population, as well as the probabilities they will choose each action with. We estimate the optimal values of the parameters through a maximum likelihood estimation. Namely, we first calculate the probability of observing the given set of choices from a given player conditional on his type, which will be simply the product of probabilities of observing each of the choices. Assuming that each player is of a given type and that this remains unchanged throughout the experiment, the probability of making the actual choices conditional on being of type k is simply the product of probabilities of making each individual choice.

$$PR_i^k(\gamma_1, \mu_1, \gamma_2, \alpha, \beta) = \prod_t p_{it}^k$$

Then, the ex–ante likelihood of this player's choices is

$$L_i(\gamma, \mu_1, \gamma_2, a_0, a_1, \alpha, \beta) = a_0 PR_i^0 + a_1 PR_i^1 + (1 - a_0 - a_1) PR_i^2$$

where a_0 and a_1 are the proportions of players of Level–0 and Level–1 in the population. Finally, the likelihood used for the estimation is

$$L = \sum_i log\left[L_i(\gamma, \mu_1, \gamma_2, a_0, a_1, \alpha, \beta)\right]$$

Note that except of the five parameters of the model, we would like to estimate also the effect (if it exists) of the two levels of virtual observability. In fact, the existence of both α and β in our model allows us to consider the possibility of different levels of depth in the reasoning associated with virtual observability. In contrast to the idea of bounded rationality itself, low values of α and β do not imply necessarily the existence of bounds in the levels of rationality of the agents, but rather bounds on the extent to which they perceive their opponents as being able to anticipate some choices they are not able to observe.

4. The Experiment

4.1. Experimental Design

The experiment took place at the Laboratory for Experimental Economics at the University of Cyprus (UCY LExEcon). A total of 96 subjects were recruited in 8 sessions, with 12 subjects in each session. Average total payment was approximately 13.7 euros and the experiment lasted about 80 minutes.[6] The experiment lasted for 50 rounds, prior to which there were 4 practice rounds to help the subjects familiarize with the experimental environment.[7] The experiment was designed on z-Tree (Fischbacher, 2007).[8]

There were two treatments, each corresponding to one of the two versions of the Selten's Horse game described above. Subjects participating in a session were playing only one of the treatments in all rounds. The subjects of each session were split into two completely separate groups of six (although this was not mentioned). This allowed us to gather two completely independent sets of observations, which could be useful for subsequent statistical tests. In each round the subjects of each subgroup were split into groups of three people, which would be the three players that would interact in that round.

The subjects were also assigned one of the three roles in the game, which means that all subjects ended up playing all different roles several times during the experiment. Changing the roles of the players

[6] Final earnings were determined by the sum of subjects' earnings throughout the 50 rounds of the experiment. Practice rounds were not taken into account for determining final earnings. Payment included 5 euros as participation fee.

[7] Due to a technical problem in round 50 of one session, we have excluded this round from our analysis for all sessions.

[8] Complete instructions can be found in the Appendix.

bears the advantage of allowing the players to experience all roles, thus understanding better the strategic implications of different strategies of different roles. However, we are aware that it might speed down learning and convergence of individual strategies, which in turn might lead to more mixed results.[9] Moreover, it allows us to consider choices in all three roles to establish the likelihood of a player being of each type in our structural model.

All payoffs were multiplied by 10 compared to the original game in order to make sure that the subjects would perceive payoff differences as sufficiently important to incentivize them to exert the required effort to understand the consequences of their different choices.

We conducted the experiment using the *strategy method*. Recall, that in the original game (which corresponds to experimental Treatment 1), P2's decision would matter only in case P1 chose action C and P3's decision would matter only if at least one of the other two players chose action D or d, respectively. In the current design, the strategy method means that P2 and P3 were asked to make a choice conditional on the fact that their choices would matter. More specifically, P2 was asked to choose between her two available choices, for the case that P1 has chosen action C. Similarly, P3 was asked to choose between her two available choices, for the case that either P1 has chosen action D, or P2 has chosen action d. Treatment 2 was conducted in an exactly analogous manner.[10]

We are aware that the strategy method may be controversial, but with the specific game being studied a direct response approach might have led to a very limited number of observations as in both treatments of the game there are equilibria where one of the players does not have an active payoff–relevant choice. Nevertheless, we acknowledge that this choice might affect our results for which the sequential nature of the game is important. As an attempt to reinforce the sequential nature of the game, we imposed decisions to be made sequentially during each round. That is, P1 was choosing first, all other players had to wait, then P2 was asked to choose and although she could not observe whether P1 had actually chosen C, she knew that P1 had already chosen. And, for all that time P3 had to wait. Finally P3 was asked to choose what he would do if the choices of the earlier players lead to his information set (and of course at the time of the choice, he does what transpired with regard to the earlier players' choices).

At the end of each round the subjects were informed about the choices of all three players of their group and their payoff for that round. This is another modelling choice that comes at a cost, as the perceived

[9]In Neugebauer, Sadrieh, and Selten (2022) subjects kept the same role for the entire session. This is one among several differencesbetween the two experimental designs. For instance, in Neugebauer, Sadrieh, and Selten (2022) subjects learn about the other players by observing them from their own role. It is an interesting question as to whether subjects learn quicker by observing the other players (Neugebauer, Sadrieh, and Selten, 2022, specification) specification) or by being those other players (our specification).

[10] In the Appendix, we include also the exact framing of the questions to which subjects had to answer depending on the role they had in each round, as this might have an impact in the thought process they follow in the current setup.

feedback imposes dependency across the choices of different subjects of the same group.

4.2. Testable Hypotheses

In the following section, we summarize the theoretical predictions that we aim to test with our experiment. First of all, due to the existence of the continuums of mixed strategy Nash equilibria, we will consider equilibrium outcomes CcR and CcL together, denoting them by $Cc(\cdot)$, and similarly for DcL and DdL, denoting them by $D(\cdot)L$. Given this, if sequential rationality were to be supported by equilibrium play then at least we should expect to observe $Cc(\cdot)$ being played more often than $D(\cdot)L$ in both treatments, which is our first testable hypothesis.

Hypothesis 1a: Equilibrium outcomes $Cc(\cdot)$ are observed more often than $D(\cdot)L$ in both treatments.

However, one might argue that looking exclusively at equilibrium outcomes might not be able to capture accurately the tendency of some types of players towards one or the other strategy, therefore we also perform comparisons on strategies used by different types of players.

Hypothesis 1b: P1 chooses action C more often than D, P3 chooses action R more often than L and P2 chooses action c equally often in both treatments.

The next questions is whether IRE can indeed act as a more accurate predictor of equilibrium behavior, instead of behaviors dictated by refinements based on sequential rationality. If this is the case, then we could identify some differences across treatments. Recall that, the only difference between the two treatments is that in Treatment 2 P3 plays before the other two players. Proponents of normal form invariance[11] will argue that this should not have an effect on the choices of the players, given that no player is informed about the choice of P3 before the end of the game. However, certain behavioral aspects associated with the effect of the order of play might lead to different outcomes between the two games, as strong order of play effects have been seen in the literature.

Hypothesis 2a: Equilibrium DcL is observed more often in Treatment 1 than in Treatment 2, whereas equilibrium CcR is observed more often in Treatment 2 than in Treatment 1.

[11] Thompson (1952) and Dalkey (1953) developed a set of conditions under which equilibrium of extensive form games will be invariant. For years these have been held by the literature as very basic. Among these was normal form invariance, which stated that games with the same normal form should have the same equilibria. Recent experimental literature has conclusively shown that normal form invariance breaks down.

Confirmation of this hypothesis will naturally imply a breakdown of normal form invariance, since with normal form invariance players should view both treatments as identical. Similarly to the previous argument, we would like to test this hypothesis considering also the frequencies of individual choices.

Hypothesis 2b: P1 chooses action D (resp. C) more often than C (resp. D) in Treatment 1 (resp. 2) then in Treatment 2 (resp. 1), P3 chooses action L (resp. R) more often than R (resp. L) in Treatment 1 (resp. 2) than in Treatment 2 (resp. 1) and P2 chooses action c equally often in both treatments.

Yet, the importance of confirmation of the previous hypotheses lies on the presumption that predicted equilibrium outcomes will be observed sufficiently often. However, as we already saw in our theoretical analysis strategies may well differ from those prescribed by Nash equilibria if subjects either behave on a boundedly rational way, or they are affected by virtual observability in a probabilistic way, or both. Hence, even if the answers to the previous hypotheses do not provide conclusive results, we would like to affirm a potential impact of virtual observability based on our alternative model.

Hypothesis 2c: α and β are different than zero in both treatments.

Note that the maximum likelihood estimation will provide us a point estimate from which we cannot make inference about its significance. For this reason, we will run a simple bootstrap on our data and generate a distribution of the values of these parameters, which we then use to make inference.

If our bounded rationality approach is valid, then we expect to observe off–equilibrium strategy profiles rather often. Of these we are mostly interested in the profile CdL, which is the one that gives the maximum payoff to P1 and P2 and nothing to P3. Which of the two variants of the game makes it more likely is ambiguous. If the order of play is unimportant then we should not expect to observe it more frequently in one of the treatments; but what happens if the order does matter? On one hand, if we consider players of both treatments to have similar distributions of cognitive hierarchies then the outcome would depend on the distribution of the sophistication levels of the subjects. In more naive populations CdL would be observed more in Treatment 2, whereas in more sophisticated populations it would be observed more often in Treatment 1 (see indicatively Table 1). The intuition here lies mostly on the behavior of P3 in Treatment 2 and is related to the ide of first–mover (dis)advantage. More specifically, players who move last can anticipate the tendency of the first–mover to lead the game towards his most favorable outcome and may be able to exploit his commitment by reaching systematically off–equilibrium action profiles that give them higher payoff, if the first mover does not anticipate this because she has not thought out the process completely. This is possible if P3 is sufficiently naive. Higher level of sophistication would allow him to anticipate this behavior and counteract by choosing R more often. The level of sophistication is not only related to cognitive ability, but also to the intensity of the perception of virtual observability. That is, even a rather naive P3 with very high β can anticipate that subsequent player will be able to observe his commitment on L and exploit it, making him more likely to choose R. Hence, the first hypothesis should be related to the relative frequency of CdL in the two treatments.

Hypothesis 3a: Profile CdL is played in equilibrium more often in Treatment 2 than in Treatment 1.

Moreover, it is apparent that a tendency to drive P3 towards CdL would come at the expense of more rounds in which P1 and P2 would end up at CdR with zero profits. The question is whether this strategy pays off, in that it leads to higher profits in Treatment 2 compared to Treatment 1 and more importantly whether it yields higher payoffs than the payoff–dominant equilibrium.

Hypothesis 3b: P1 obtains higher profits in Treatment 2 than in Treatment 1 and higher average profits than 30 in both treatments.

5. Results

In what follows, we present the major results from our experiment. Recall that our design allowed us to have eight fully independent data points for each treatment, corresponding to the set of observations from the same community. Therefore, all our econometric tests are done at the community level by pooling our observations in the way that is appropriate for each hypothesis. In addition to this, for each community of Treatment 1, there was exactly one community of Treatment 2 with the same order of assigned roles. This allows us to pair observations from the two treatments, thus use the respective tests for paired observations. For the estimation of the structural model we use individual data at the treatment level. We use individual data also for some regressions, in which we take into account the correlation that may exist between observations gathered from the same community.

Result 1a: Equilibria $D(\cdot)L$ are played significantly more often than $Cc(\cdot)$ in both treatments.

This is the first benchmark result of the experiment. Hypothesis 1a of sequential rationality is strongly rejected. The result can be readily observed in Figures 3, 4 and Table 2. More formally, we compare the relative frequencies for each community of each treatment, with the null hypothesis being that mean/median/distribution (depending on the test) are the same. We find strongly significant and very similar results for both treatments. The Wilcoxon signed–rank test rejects the null at 5% level ($z = -2.100$, p–value 0.0357 for Treatment 1 and $z = -2.240$, p–value 0.0251 for Treatment 2) and similarly for the t–test ($t = -2.8741$, p–value 0.0119 for Treatment 1 and $t = -3.6001$, p–value 0.0044 for Treatment 2) and for the sign test (p–value 0.0352 for both treatments).

Overall, it is striking how rarely we observe CcR as an outcome in either treatment, suggesting that sequential rationality is not prominent is this game.

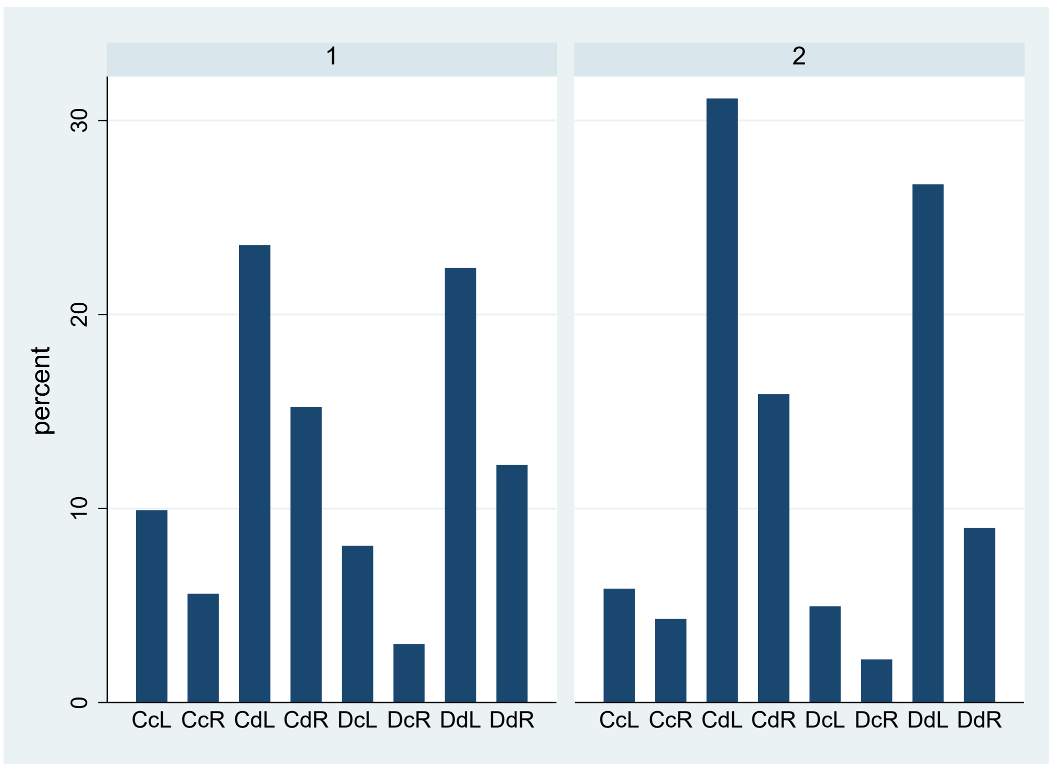


Figure 3: Frequency of observations for each strategy profile, per treatment.

Result 1b/2b: P1 chooses action D slightly more often in both treatments. P2 chooses strategy d much more often in both treatments. P3 chooses action L much more often in both treatments.

Neither Hypothesis 1b nor 2b is supported by the data. In fact, we observe P2 and P3 playing very often actions d and L, respectively, whereas P1 use both strategies almost equally often. See Figure 5 for a complete picture of the evolution of choices of all player types during the game. Although the graphical evidence is overwhelming, we verify its statistical significance through the relevant tests, where we compare at the community level the observed frequencies first with 1/2. For P1 none of the performed tests rejects

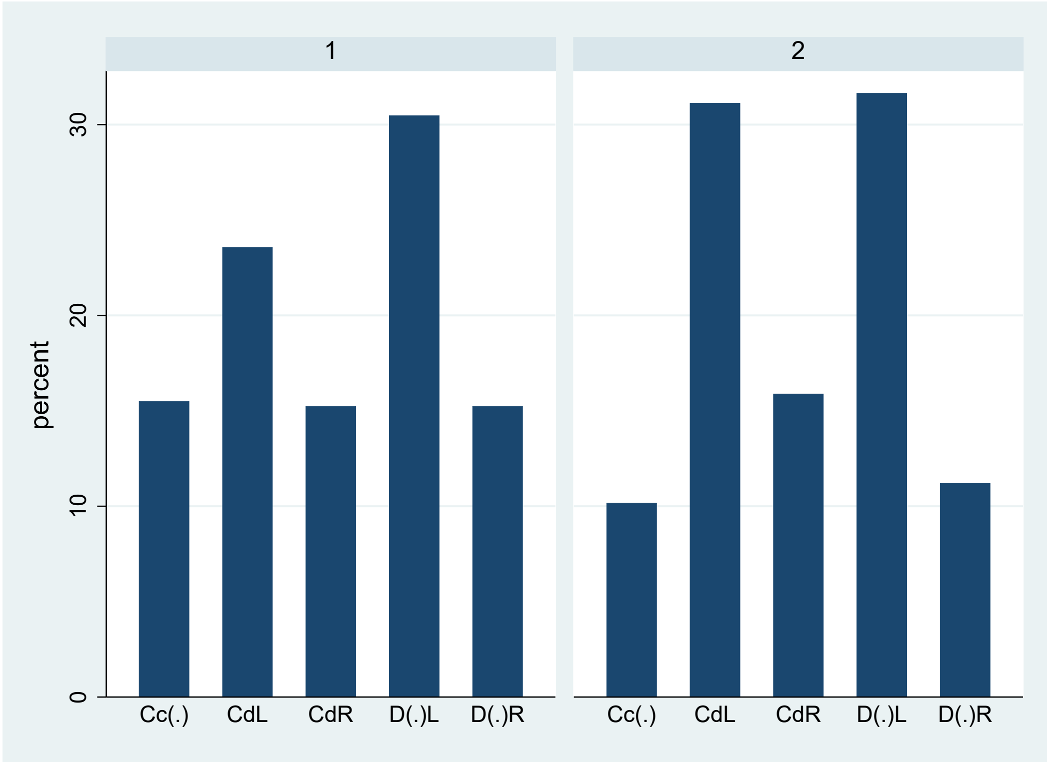


Figure 4: Frequency of observations of essentially different strategy profile, per treatment.

the hypothesis that each strategy is used with equal probability for either treatment[12]. For P2 all tests reject the null hypothesis for both treatments.[13] In fact, the same results hold if we consider a null hypothesis suggesting that d is chosen with probability 3/5. Finally, for P3 almost all tests again reject the null hypothesis of equal probabilities for both treatments. [14]

Result 2a: Equilibrium DcL is observed more often in Treatment 1 than in Treatment 2, whereas CcR is not observed more often in Treatment 2 than in Treatment 1.

The result is apparent from Table 2 and is further supported by a Wilcoxon rank–sum test performed on the occurrences of each of the two equilibria in each treatment at individual level (p–value 0.0022 for DcL and 0.2930 for CcR). The result is the same if we perform a Fischer's exact test to compare the medians (p–value 0.025 for DcL and 0.350 for CcR). However, due to the small sample size, we do not have sufficient

[12]Treatment 1: Wilcoxon, $z = 1.054$, p–value 0.2918. T–test, $t = 1.2179$, p–value 0.1314. Sign test, p–value 0.2266. Treatment 2: Wilcoxon, $z = 1.407$, p–value 0.1594. T–test, $t = 1.6597$, p–value 0.0705. Sign test, p–value 0.1445.

[13]Treatment 1: Wilcoxon, $z = -2.380$, p–value 0.0173. T–test, $t = -4.7266$, p–value 0.0011. Sign test, p–value 0.0352. Treatment 2: Wilcoxon, $z = -2.521$, p–value 0.0117. T–test, $t = -6.4605$, p–value 0.0002. Sign test, p–value 0.0039.

[14]Treatment 1: Wilcoxon, $z = -2.103$, p–value 0.0355. T–test, $t = -2.7307$, p–value 0.0147. Sign test, p–value 0.1445. Treatment 2: Wilcoxon, $z = -2.380$, p–value 0.0173. T–test, $t = -5.3645$, p–value 0.0005. Sign test, p–value 0.0352.

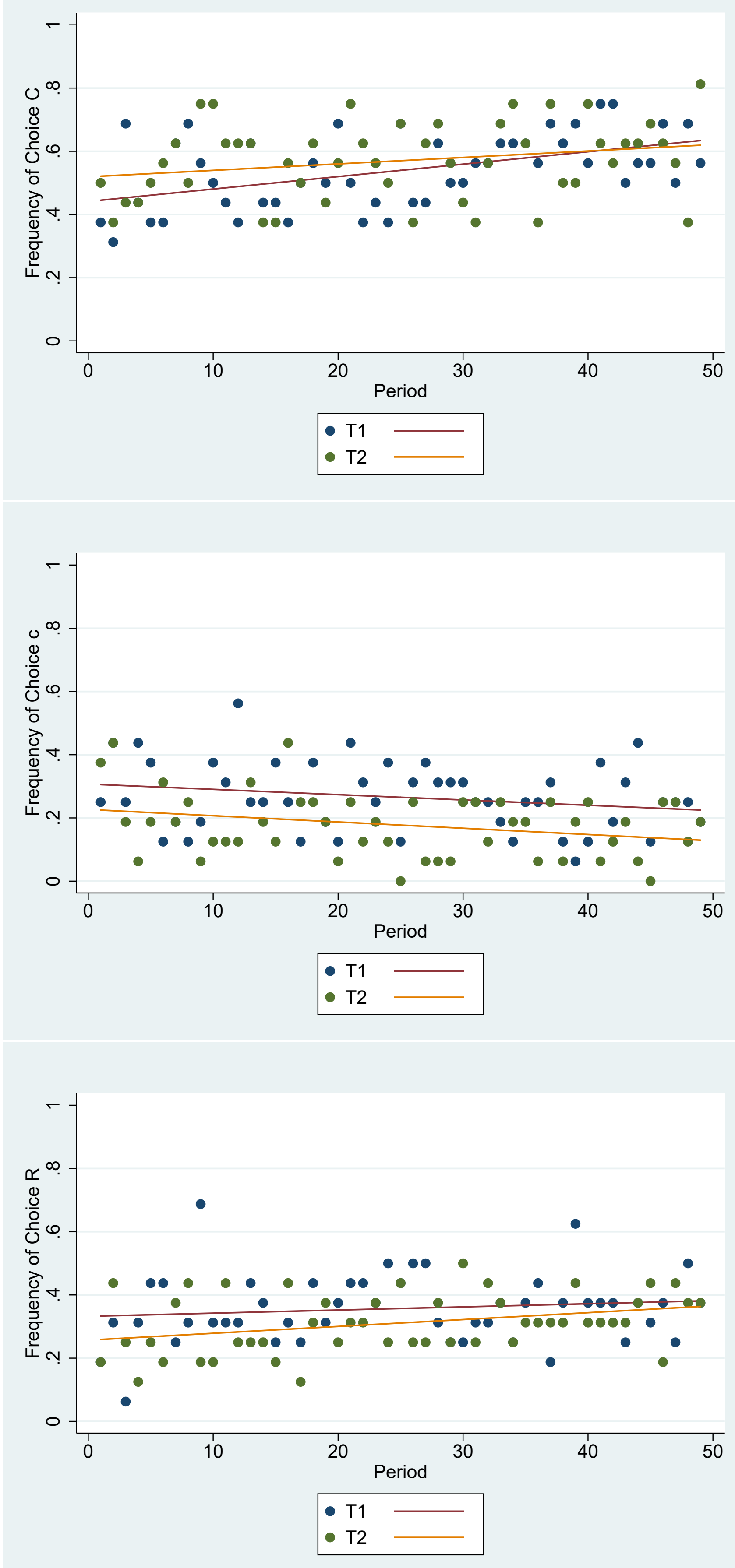


Figure 5: Evolution of choices of the three players (P1, P2 and P3 from top to bottom), per treatment.

Strategy Profile	Treatment 1	Treatment 2
	784 obs.	784 obs.
DcL	63	40
CcR	43	34
$D(\cdot)L$	241	249
$Cc(\cdot)$	120	82

Table 2: Frequencies of equilibrium observations

power to support the result at a community level (p–values of the Wilcoxon signed–rank test are 0.1609 for DcL and 0.29 for CcR, for the sign test are 0.1445 and 0.2266 respectively and for the t–test[15] are 0.0892 and 0.2429 respectively.)

We could argue that the result provides some evidence in support of IRE, however both equilibria are very seldom played in both treatments, mainly due to the strong tendency of P2 in favor of action d. The result is similar for Hypothesis 2b, which can be answered through the findings of Result 1b/2b, and we do not find evidence to support it.

Result 2b: This result is not supported by the data; we have already discussed this under Result 1b.

The previous findings suggest that our attempt to understand the impact (if it exists) of virtual observability on subjects' choices should be focused on the results of the structural model. Interestingly, taking a look at the occurrences of different strategy profiles (see Figure 3), their frequencies seem to be in line with the projection of a Level–k model with predominantly players with rather naive (recall the simple exercise presented in Table 1). This is quite encouraging regarding the potential adequacy of the proposed model to capture the actual behavior of the experimental subjects.

Result 2c: Parameter α is most likely different than zero in Treatment 2 and likely to be different than zero in Treatment 1, while β is likely to be different than zero in Treatment 2 and most likely equal to zero in Treatment 1.

Table 3 presents the results of three different specifications of the model. One in which both levels of

[15] Of course, throughout our analysis, the t–test's results should be taken into account with even more caution, given the parametric assumptions that the test makes regarding the distribution of data and the small sample size when looking at independent communities.

virtual observability are allowed, another in which only removal of all information sets is allowed ($\alpha = 0$) and one in which no virtual observability is present $\alpha = \beta = 0$. There are several interesting takeaways from this estimation. First and most important is that the ability/ willingness to choose taking into account virtual observability is limited to one step ahead. That is, the subjects do not consider possible that information sets are removed for players who are far from them. That is important, because it suggests that, in Treatment 2, P3 does not consider as possible that P2 might be able to anticipate his action. This observation is reinforced even further by looking at the specification that contains only $\beta \geq 0$. In this specification the additional parameter does not offer anything, as the prediction is exactly the same as in the case where virtual observability is completely ruled out. On the contrary, when allowing for partial observability, in the sense that choices are perceived to be visible only by immediate successors, we observe this to be something that subjects take into account in their decisions.

Overall, we find that indeed virtual observability seems to play a role in the players' choices, but in a more limited and possibly naive way compared to what equibrium predictions would require.

There is a number of other interesting features in the results, as for instance the fact that even Level–2 players when having the role of P3 appear to choose L around half of the times. This is because, despite the fact that sophisticated players of types P1 and P2 might try to guide P3 towards the bad outcome, the existence of naive players makes the choice of L sufficiently attractive. A last observation is that the estimated parameters suggest that the game of Treatment 2 possesses a higher level of complexity.

The estimated parameters are not sufficient to answer the question of our Hypothesis 2c. In order to do that, we have performed a simple bootstrap on the original samples. More specifically, for each treatment, we have generated 1000 random resamples chosen from the original sample with replacement choosing observations from the same treatment. One observation is considered to be the distribution of choices of one individual over the all rounds. We have then used this new data set to estimate again the values of all parameters and have constructed a distribution of the many estimated parameters. The result of the bootstrap for α and β are presented in Figure 6, with some additional statistics on Table 4. Although the lower bounds of the confidence intervals are almost at zero, most of the estimated parameters are sufficiently far away from this lower bound. The last result that we obtain from this empirical exercise has to do with the observed sophistication of players in different treatments. Arguably, Treatment 2 seems to be slightly more complicated for the players, mainly for P3, and particularly in an environment that allows for virtual observability. This result seems to be verified by the data as well, in the sense that we find a lower proportion of players of Level–2 in Treatment 2 compared to Treatment 1.

	$\alpha, \beta \geq 0$							
	Treatment 1				Treatment 2			
Level	p	r	q	Proportions	p	r	q	Proportions
Level–0	0.500	0.500	0.500	0.325	0.500	0.500	0.500	0.158
Level–1	0.500	0.868	0.842	0.268	0.500	0.841	0.835	0.536
Level–2	0.307	0.927	0.643	0.407	0.205	0.983	0.578	0.306
Probabilities	0.421	0.773	0.650		0.410	0.831	0.704	
	α	β	L		α	β	L	
	0.110	0.000	-1407.290		0.179	0.000	-1267.053	

	$\alpha = 0$							
	Treatment 1				Treatment 2			
Level	p	r	q	Proportions	p	r	q	Proportions
Level–0	0.500	0.500	0.500	0.297	0.500	0.500	0.500	0.164
Level–1	0.500	0.827	0.764	0.390	0.500	0.868	0.803	0.534
Level–2	0.287	0.938	0.612	0.314	0.204	0.983	0.585	0.301
Probabilities	0.433	0.765	0.638		0.411	0.842	0.688	
		β	L			β	L	
		0.000	-1409.560			0.001	-1270.066	

	$\alpha = \beta = 0$							
	Treatment 1				Treatment 2			
Level	p	r	q	Proportions	p	r	q	Proportions
Level–0	0.500	0.500	0.500	0.297	0.500	0.500	0.500	0.160
Level–1	0.500	0.827	0.764	0.390	0.500	0.862	0.798	0.513
Level–2	0.287	0.938	0.612	0.313	0.222	0.977	0.652	0.327
Probabilities	0.433	0.765	0.638		0.409	0.842	0.703	
			L				L	
			-1409.560				-1272.044	

Table 3: Estimation results: (Top) the more general model with $\alpha, \beta \geq 0$, (Middle) restricting to $\alpha = 0$, (Bottom) No virtual observability $\alpha = \beta = 0$.

Result 2d: The proportion of Level–2 players is lower in Treatment 2 than in Treatment 1.

Evidence supporting the result are presented in Figure 6 and Table 4 and are verified also by the relevant tests, which all strongly reject the null hypothesis of equality of the distributions (Mann–Whitney), medians (Mood's median test), or means (t–test).[16]

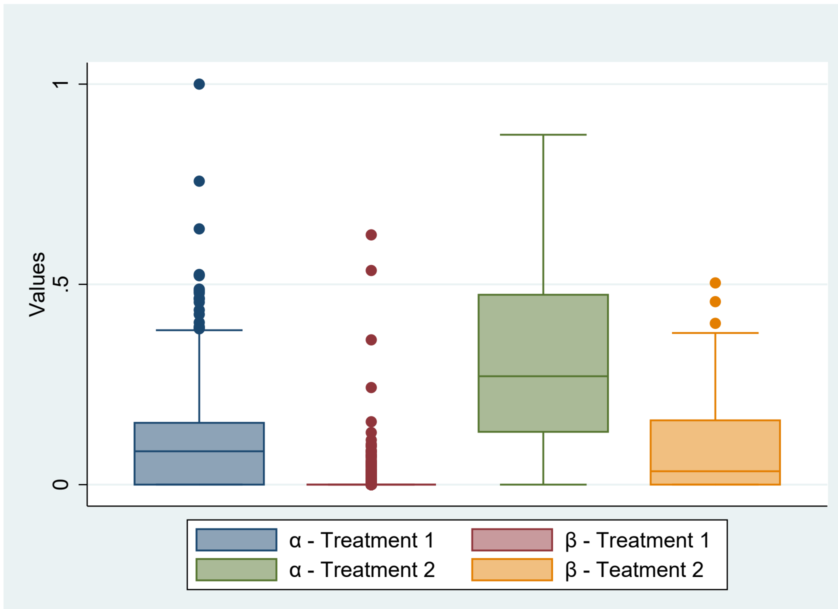


Figure 6: The boxplot of the parameters α and β in each treatment estimated from bootstrapped data.

Parameter	Treatment	Mean	Median	St. Dev.	95% Conf. Interval	$P(> 0.001)$
α	T1	0.0916	0.0832	0.1008	(0, 0.2811)	0.627
	T2	0.3019	0.27	0.2016	(0,0.7009)	0.944
β	T1	0.0047	0	0.0323	(0,0.0543)	0.074
	T2	0.0828	0.0334	0.0966	(0,0.2910)	0.673
Level–2	T1	0.3904	0.3978	0.1541	(0.0052,0.6786)	
	T2	0.2704	0.2667	0.0840	(0.1257,0.4533)	

Table 4: Descriptive statistics of bootsrapped data for parameters α and β, as well as the proportions of Level–2 players in both treatments.

Result 3a: Strategy profile CdL is played in equilibrium more often in Treatment 2 than in Treatment 1.

[16]Here we cannot used tests for paired data. Results of tests are the following: Mann–Whitney, $z = 21.364$, $p = 0.000$. T–test, $t = 21.6290$, $p = 0.000$. Median test: $\chi = 388.9620$, $p = 0.000$.

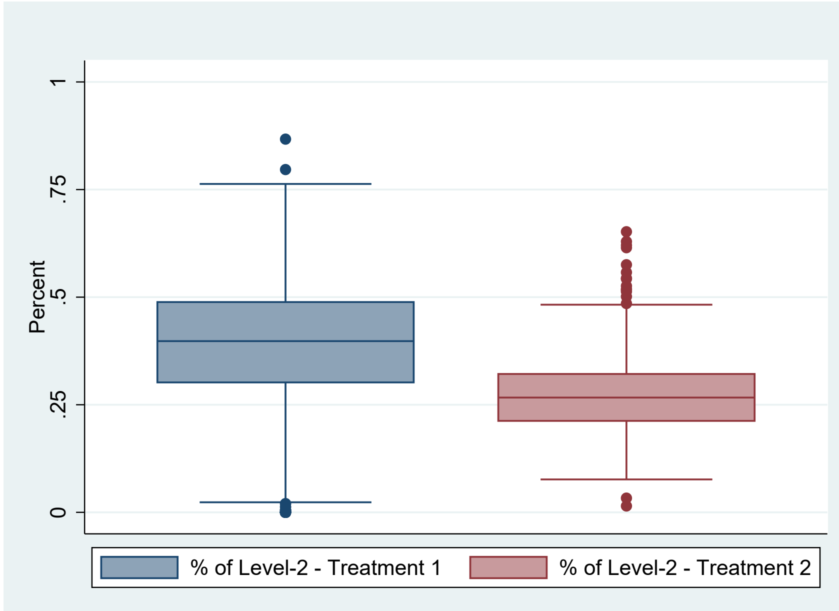


Figure 7: The boxplot of the proportions of Level–2 players in each treatment estimated from bootstrapped data.

We consider the frequency of appearance of the specific strategy profile at each community of each treatment over the course of all rounds. The t–test ($t = -2.0996$, p–value 0.0370) and the sign test (p–value 0.0352) suggest that the differences of the of the mean and median are statistically significant. The Wilcoxon signed–rank test yields a more ambiguous result ($z = -1.540$, p–value 0.1235), probably due to the limited sample, as among the eight pairs of data points seven are in the predicted direction.

It is still interesting to observe that this tendency is reduced over rounds, which could be accounted to learning, but its explanation seems to be slightly more convoluted than that. In fact, if we split behavior into early and late rounds (Figure 8), we can see a similar frequency of occurrences of CdL in Treatment 2, which is however associated with an increase in the occurrences of CdR. This suggests that subjects playing as P3 seem to become more aware of the strategy followed by the other two players and to respond to that. This reaction can indeed by accounted to learning. On the contrary, in Treatment 1, we observe a slightly opposite effect. In the beginning all players coordinate strongly towards the payoff–dominant equilibrium, but as the rounds progress, subjects who get to play as P1 become sufficiently confident that the corresponding P3 will choose L, in which case they become tempted to play C and achieve their maximum payoff through profile CdL. This reduces slightly the difference between the frequency of CdL observed in the two treatments, which however remains significant.

Yet, is such behavior beneficial for P1 and P2? The answer is once again relative: On the one hand, a successful coordination to the payoff–dominant equilibrium would have allowed P1 and P2 to obtain higher profits than they actually do. On the other hand, when comparing among treatments, it seems that the

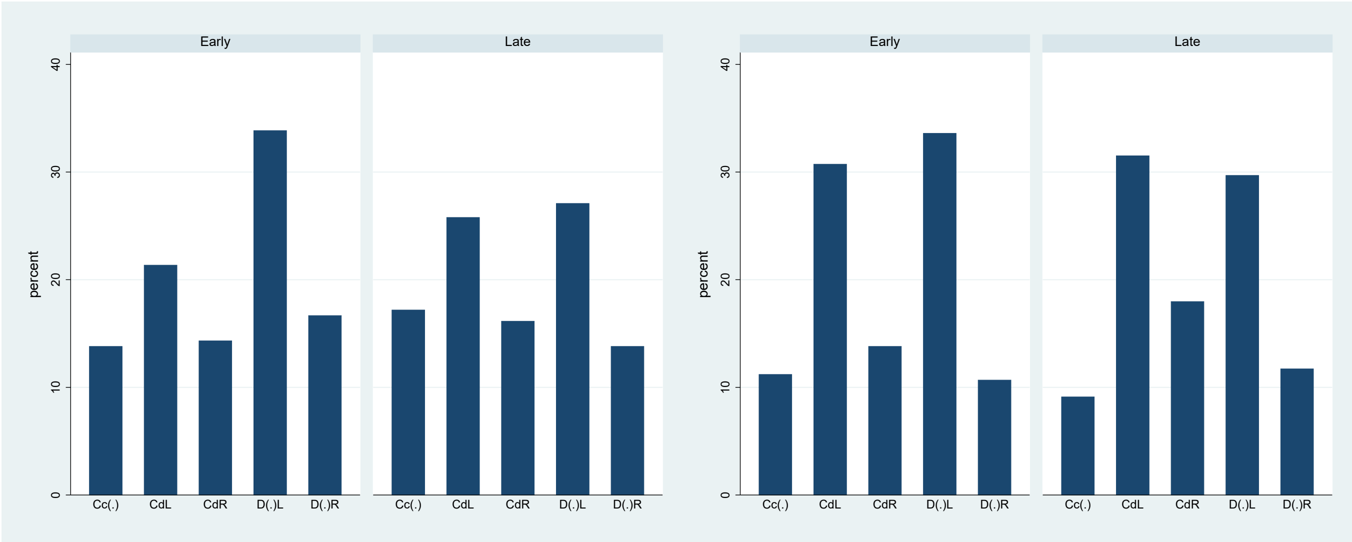


Figure 8: Frequency of observations of essentially different equilibria, split between early (< 25) and late (> 25) rounds, per treatment. Frequencies appear as percentages of observations in given round span and treatment.

anticipation of the tendency of P3 to move towards her most preferred equilibrium, allows P1 and P2 to enjoy higher profits. Formally,

Result 3b: P1 obtains higher profits in Treatment 2 than in Treatment 1.

Figure 9 shows the average per round profits of P1 and P3 in the different groups of each treatment in increasing relative order. On top of the suggestive graphical evidence, comparing the profit distributions via a Wilcoxon signed–rank test we can reject the null hypothesis of equal profits at the 5% significance level ($z = -2.100$ and p–value 0.0357) and similarly when comparing their means via t–test ($t = 0.0191$ and p–value 0.0382). The results of the sign test on the comparison of medians is less conclusive (p–value is 0.1445 for the one–sided test). The profits of P3 are almost indistinguishable between the two treatments.

Interestingly, this observation regarding P1's profits seems to be more prominent in the first periods. If we perform the same analysis separately for the first rounds (≤ 25) and the last ones, then the result holds for the former, but not for the latter. More specifically, for the first 25 rounds all three tests yield results that are statistically significant: Wilcoxon signed–rank test ($z = -2.176$, p–value 0.0296), t–test ($t = -3.4281$, p–value 0.0055) and sign test (p–value 0.0625). On the contrary, for the last 25 rounds the effect seems to be reduced if not disappear completely: Wilcoxon signed–rank test ($z = -1.120$, p–value 0..2626), t–test ($t = -1.3166$, p–value 0.1147) and sign test (p–value 0.3633). This observation is in line with our previous findings regarding the second–mover advantage. As it seems, the ability of P1 and P2 to

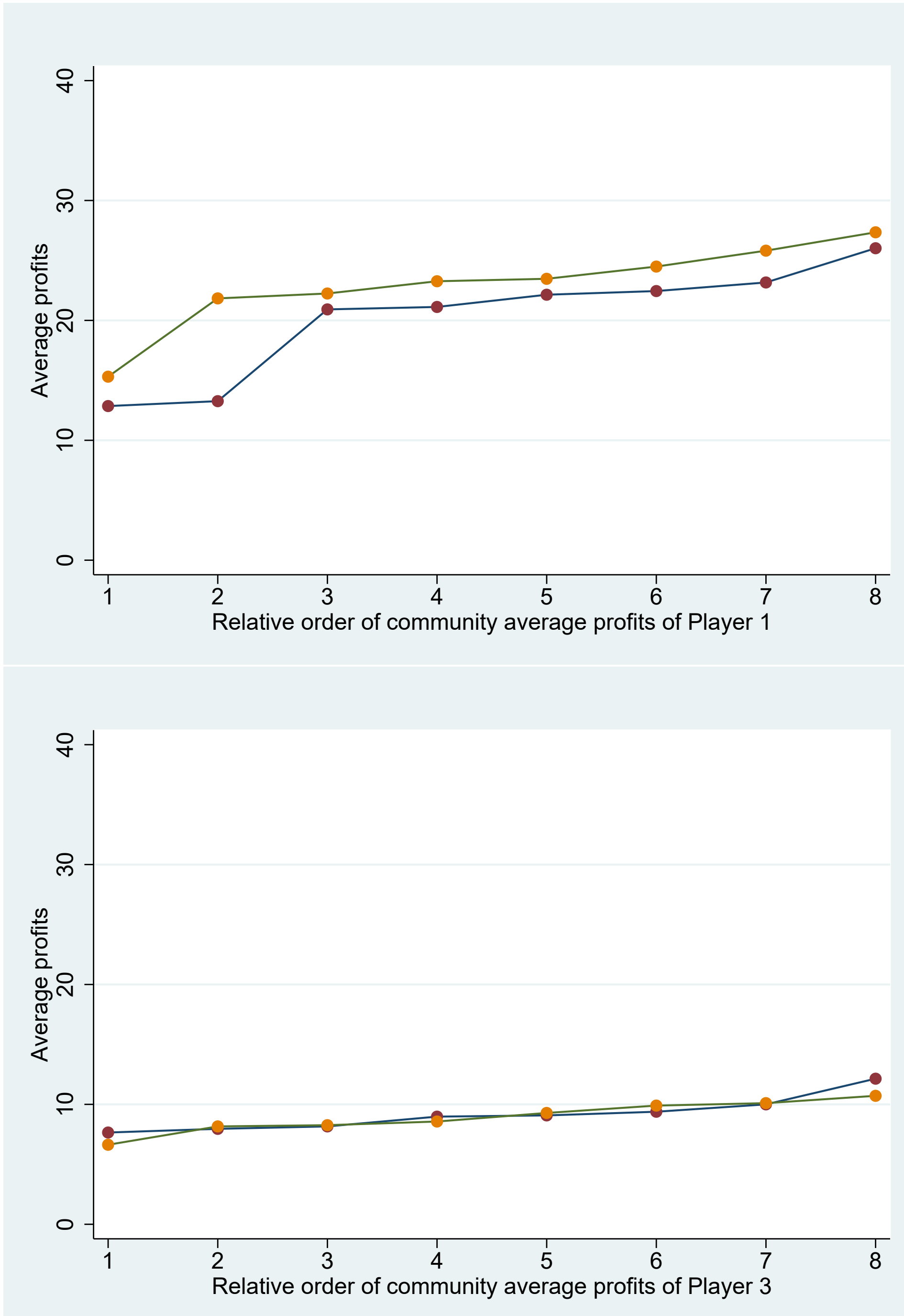


Figure 9: Average per round profits of Players 1 and 3 respectively over the course of the whole experiment. Orange dots correspond to Treatment 2, red dots correspond to Treatment 1.

induce P3 to move towards their most preferred outcome is more prominent in the initial rounds, as it then can be often anticipated.

Result 3c: P1 obtains on average lower profits than 30 in both Treatments.

The result is already evident from Figure 9, as all groups receive an average profit lower than 30 and is the null of equality in distribution/mean/median is rejected by the relevant tests.[17]

Impact of individual choices on profits

So far, we have not focused on the individual performance of subjects, but only on the different types of players. For this, we also perform two sets of regressions that relate individual choices to profits. All regressions are taking care of dependency between subjects of the same community. In Table 6 we look at the relation between the frequency of individual choices and average profits. It turns out that avoiding action c when having the role of P2 is the only consistently important choice for the subjects' profits. Interestingly, when focusing on Treatment 2, playing C more often is associated with higher profits, which means that attempting to exploit the anticipated tendency of first–mover (P3) towards her preferred equilibrium was indeed a profitable strategy in this treatment. One possible issue with this set of regressions is that average profits are affected by rounds in which subjects were playing different roles, hence a strong effect from some rounds might dominate the results. For this reason, we also perform a second set of regressions (see Table 5) in which we consider average profits conditional on the role the subject played (i.e. average profits of subject i when being Player j) and we look at the correlation of these with the choices the subject made during these rounds. Interestingly, P3 playing L (mainly) in Treatment 2 seems to have had a detrimental effect on profits. This suggests that, despite the possibility of being exploited at times, the fact that P1 players did not play C too often, was sufficient to make L a profitable choice for P3.

6. Discussion: Alternative behavioral rules

In this last part, we consider some alternative behavioral explanations that could give rise to the data we observe. First, we have estimated a Quantal Response Equilibrium (QRE) model (McKelvey and Palfrey, 1995; McKelvey and Plafrey, 1998), in which agents follow fully mixed strategies, but not completely random. In QRE the probability of choosing a strategy increases in the expected payoff of that strategy. More specifi-

[17] For Treatment 1: Wilcoxon, $z = -2.521$, $p = 0.0117$. Signtest, $p = 0.0039$. T–test, $t = -5.8705$, $p = 0.0003$. For Treatment 2: Wilcoxon, $z = -2.521$, $p = 0.0117$. Signtest, $p = 0.0039$. T–test: $t = -5.5288$, $p = 0.0004$.

	Average Profits			
	T1 & T2	T1 & T2	T1	T2
P1_C	0.771	0.737	-1.345	2.541*
	(1.313)	(1.327)	(2.529)	(1.029)
P2_c	-6.029***	-5.696***	-5.522**	-6.669**
	(1.139)	(1.018)	(1.542)	(1.368)
P3_R	-2.285	-2.140	-2.654	-1.243
	(1.267)	(1.231)	(1.953)	(1.105)
Treatment		1.171		
		(1.102)		
Constant	19.13***	17.27***	19.72***	18.49***
	(0.914)	(1.732)	(1.492)	(0.885)
Cluster SE	Y	Y	Y	Y
N	96	96	48	48

Standard errors in parentheses, clustered at Community level,

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5: Regression of average profits on frequencies of choices.

cally, for an agent i the probability of choosing action k is given by a logistic function: $P_{ik} = \frac{exp(\lambda EU_{ik})}{\sum\limits_{j \in J_i} exp(\lambda EU_{ij})}$. The level of randomness depends on a single parameter λ, where $\lambda = 0$ signifies uniform mixing over actions and $\lambda = \infty$ signifies pure rationality. Note that, this method can predict actual equilibrium behavior. Recall that a similar mechanism was employed in our structural model, as we considered probabilistic choices based on the relative payoffs of the two choices. The main difference is that here agents do not choose according to some arbitrarily formed beliefs regarding the sophistication of their opponents, but best respond to their actual play. This of course requires a certain level of sophistication, higher than the one we considered, but part of it can be captured by the value of parameter λ.

One way to estimate the parameter λ is to use the empirical frequencies at which each action was used by ones opponents in order to calculate the expected utilities and then use maximum likelihood estimation (see Goeree, Holt, and Palfrey, 2016, for details). We do so here both at a community and at a treatment level and we find results that are quite close to the actual observations (see Tables 7 and 8), despite the fact that we estimate a single parameter.

	Avg. Profits when P1			Avg. Profits when P2			Avg. Profits when P3		
	T1 & T2	T1	T2	T1 & T2	T1	T2	T1 & T2	T1	T2
P1_C	2.154	0.628	3.483						
	(2.629)	(5.445)	(1.649)						
P2_c				-13.51***	-12.66***	-14.65***			
				(1.304)	(1.741)	(2.113)			
P3_R							-4.960**	-4.286	-5.540**
							(1.271)	(2.108)	(1.528)
Treatment	2.777			1.484			-0.424		
	(2.134)			(1.596)			(0.664)		
Constant	16.21***	19.83***	21.01***	22.30***	23.55***	25.48***	11.30***	10.64***	10.63***
	(3.001)	(2.007)	(1.097)	(2.914)	(1.462)	(0.821)	(1.209)	(1.089)	(0.925)
Cluster SE	Y	Y	Y	Y	Y	Y	Y	Y	Y
N	96	48	48	96	48	48	96	48	48

Standard errors in parentheses, clustered at Community level, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Regression of average profits conditional on type on frequencies of choices.

Nevertheless, we observe that there is a lot of updating in players' strategies during the game, which we would like to capture as well. For this reason, we look at the possibility that players follow fictitious play (see Fudenberg and Levine, 1999, for a detailed list of references), i.e. at each point in time the beliefs they have about the future play of their opponents coincide with the empirical distribution of choices they have observed in the past and they best respond to that. There is no estimation involved in this method. In the simplest case, we just calculate the best–response of each player at each round and look at how often their actual play coincides with their fictitious–play best–response. It turns out that P2 and P3 play more than 2/3 according to the best–response of fictitious play, whereas for P1 the percentage is significantly lower, to the extent that it does not seem to have very strong explanatory power for them (see Table 9).

7. Conclusion

We have conducted an experiment on the well–known game of Selten's Horse, in an attempt to test the empirical validity of equilibrium refinements that require sequential rationality, as well as the impact of the

Treatment	$\hat{\lambda}$	P_C (QRE)	P_c (QRE)	P_R (QRE)	P_C (data)	P_c (data)	P_R (data)
1	0.071	0.540	0.247	0.408	0.540	0.265	0.357
2	0.095	0.589	0.159	0.408	0.570	0.177	0.311

Table 7: Frequencies predicted by QRE vs. observed in data, per treatment

Session	Treatment	Community	$\hat{\lambda}$	P_C (QRE)	P_c (QRE)	P_R (QRE)	P_C (data)	P_c (data)	P_R (data)
1	1	1	0.074	0.544	0.223	0.452	0.633	0.255	0.327
1	1	2	0.039	0.493	0.302	0.397	0.357	0.398	0.214
2	2	1	0.125	0.661	0.094	0.414	0.592	0.092	0.296
2	2	2	0.091	0.593	0.177	0.348	0.459	0.153	0.327
3	1	1	0.099	0.602	0.174	0.355	0.490	0.143	0.357
3	1	2	0.000	0.500	0.500	0.500	0.633	0.520	0.622
4	2	1	0.135	0.652	0.107	0.570	0.765	0.112	0.357
4	2	2	0.027	0.505	0.435	0.473	0.622	0.429	0.510
5	1	1	0.099	0.603	0.147	0.388	0.541	0.153	0.306
5	1	2	0.118	0.580	0.271	0.404	0.602	0.224	0.541
6	2	1	0.127	0.681	0.078	0.416	0.592	0.071	0.265
6	2	2	0.041	0.505	0.329	0.452	0.582	0.367	0.316
7	1	1	0.051	0.515	0.261	0.416	0.500	0.316	0.245
7	1	2	0.107	0.638	0.103	0.400	0.561	0.112	0.245
8	2	1	0.089	0.609	0.111	0.432	0.592	0.143	0.163
8	2	2	0.096	0.650	0.131	0.288	0.357	0.051	0.255

Table 8: Frequencies predicted by QRE vs. observed in data, per community

order of play on behaviour. Consistent with the finding of experimental literature, we have seen that human subjects mostly do not play the equilibria predicted by these refinements. This is an indication that experimental subjects may be more inclined to use strategies that arise from more intuitive and computationally simpler thought processes. One such equilibrium concept is the Ideal Reactive Equilibrium (Sadanand, 2019) and we have found some empirical evidence that subjects use this type of thinking although in some instances the thinking was incomplete. We have also found that the simple model of Level–k thinking of Stahl and Wilson (1994) with three levels of sophistication explains the subjects' behaviour quite well.

A further modification of the model to incorporate virtual observability, provides some evidence that players who play early on might take into account the possibility that subsequent players could predict or anticipate their choices. An interesting feature that comes up is that virtual observability seems to be limited to immediate successors and not beyond. This is an issue that has so far not been explored in the

Treatment	Player 1 (/784)	Player 2 (/784)	Player 3 (/784)
1	425	568	519
2	420	645	525

Table 9: Frequencies of occasions in which a player played best–response according to fictitious play

literature as the focus there has mostly been in two–player single choice games, where could not exist multiple information sets, as in our game. Having said that, we were also able to examine the trade–off that first movers face in sequential games, between the ability to steer the game towards the most preferred equilibria and the commitment associated with moving before other players. We have found evidence suggesting that the first mover seeking advantage from his position but not recognizing the loss of control from commitment, might allow subsequent players to exploit the commitment imposed by the sequential nature of the game and reach off–equilibrium outcomes that can be beneficial for them and detrimental to the first mover.

Overall, this is a first attempt to address the validity of sequential rationality in an experimental environment and the generality of the results is limited by our specific experimental design. However, the results are sufficiently intriguing to suggest that additional work could shed more light on the processes that actual subjects follow in order to choose their strategies in complex games with multiple equilibria.

References

Amershi, A.; Sadanand, A.; Sadanand, V. *Manipulated Nash Equilibrium I: Forward Induction and thought Process Dynamics*; Faculty of Commerce and Business Administration University of British Columbia Working Paper 928; Faculty of Commerce and Business Administration, University of British Columbia: Vancouver, BC, Canada, 1985.

Amershi, A.; Sadanand, A.; Sadanand, V. *Manipulated Nash Equilibrium II: Properties*; Faculty of Commerce and Business Administration University of British Columbia Working Paper 1222a; Faculty of Commerce and Business Administration, University of British Columbia: Vancouver, BC, Canada, 1985.

Amershi, A.; Sadanand, A.; Sadanand, V. *Manipulated Nash Equilibrium III: Economic Applications and A Preliminary Experiment*; Faculty of Commerce and Business Administration, University of British Columbia Working Paper 1222b; Faculty of Commerce and Business Administration, University of British Columbia: Vancouver, BC, Canada, 1985.

Amershi, A.; Sadanand, A.; Sadanand, V. Player Importance and Forward Induction. *Economics Letters* **1992**, *38*, 291–297.

Brandts, J.; Holt, C.A. Forward Induction: Experimental Evidence from Two-Stage Games with Complete Information. In *Research in Experimental Economics*; Isaac, R.M., Ed.; Emerald Publishing Limited: Bingley, UK, 1992; *5*, 119–137.

Budescu, D.V.; Suleiman, R.; Rapoport, A. Positional and Group Size Effects in Resource Dilemmas with Uncertain Resources. *Organizational Behavior and Human Decision Processes* **1995**, *61*, 225–238.

Camerer, C.F.; Knez, M.; Weber, R. Virtual Observability in Ultimatum Bargaining and Weak Link Coordination Games. In *Proceedings of the Meetings of the Public Choice Society and Economic Science Association*, Houston, TX, USA, April 12–14, 1996.

Charness,G.; Defwenberg, M. Promises and Partnership. *Econometrica* **2006** *74(6)*, 1579–1601.

Cooper, R.; DeJong, D.V.; Forsythe, R.; Ross, T.W. Selection Criteria in Coordination Games: Some Experimental Results. *American Economic Review* **1990**, *80*, 218–233.

Cooper, R.; DeJong, D.V.; Forsythe, R.; Ross, T.W. Forward Induction in the Battle-of-the-Sexes Games. *American Economic Review* **1993**, *83*, 1303–1315.

Cooper R.; DeJong D.V.; Forsythe R.; Ross T.W. Alternative Institutions for Resolving Coordination Problems: Experimental Evidence on Forward Induction and Preplaycommunication. In *Problems of Coordination in Economic Activity*; Friedman, J.W., Ed.; Recent Economic Thought Series; Springer: Dordrecht, The Netherlands, 1994; Volume 35, Chapter 7, pp. 129–146.

Dalkey, N. Equivalence of Information Patterns and Essentially Determinate Games. In *Contributions to the Theory of Games*; Kuhn, H., Tucker, A., Eds.; Princeton University Press: Princeton, NJ, USA, 1953; Volume 2, pp. 217–244.

Fischbacher, U. zTree: Zurich toolbox for ready-made economic experiments. *Experimental Economics* **2007**, *10*, 171–178.

Fudenberg, D.; Levine, D. *The Theory of Learning in Games*, MIT Press: Cambridge, MA, USA, 1999.

Fudenberg, D.; Tirole, J. Perfect Bayesian Equilibrium and Sequential Equilibrium. *Journal of Economic Theory* **1991**, *53*, 236–260.

Gale, J.; Binmore, K.; Samuelson, L. Learning to be imperfect: The Ultimatum Game. *Games and Economic Behavior* **1995***8*, 56–90.

Gibbons, R. *Game Theory for Applied Economists*, Princeton University Press: Princeton, NJ, USA, 1992.

Goeree, J.K.; Holt, C.A.; Palfrey, T.R. *Quantal Response Equilibrium: A Stochastic Theory of Games*, Princeton University Press: Princeton, NJ, USA, 2016.

Guth, W.; Huck, S.; Rapoport, A. The limitations of the positional order effect: Can it support silent threats and non-equilibrium behavior? *Journal of Economic Behavior and Organization* **1998**, *34*, 313–325.

Guth,W. Schmittberger, R., Schwarze, B. An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior and Organization* **1982** 3, 367-388.

Guth, W. M. Kocher, More than Thirty Years of Ultimatum Bargaining Experiments: Motivations, Variations, and a Survey of the Recent Literature. *CES Info Working Paper 4380* **2013** 13, 1-36.

Hammond, P. *Beyond Normal Form Invariance: First Mover Advantage in Two-Stage Games with or without Predictable Cheap Talk*; Economic Research Papers 269784; Department of Economics, University of Warwick: Coventry, UK, 2008.

Huck, S.; Müller, W. Burning Money and (Pseudo) First-Mover Advantages: an Experimental Study on Forward Induction. *Games and Economic Behavior* **2005**, *51*, 109–127.

Kohlberg, E.; Mertens, J. On the Strategic Stability of Equilibria. *Econometrica* **1986**, *54*(5), 1003–1038.

Kreps, D.M.; Wilson, R. Sequential Equilibria. *Econometrica* **1982**, *50*(4), 864–894.

McKelvey, R.; Palfrey, T. Quantal Response Equilibria for Normal Form Games. *Games and Economic Behavior* **1995**, *10*, 6–38.

McKelvey, R.; Palfrey, T. Quantal Response Equilibria for Extensive Form Games. *Experimental Economics* **1998**, *1*, 9–41.

Muller, A.; Sadanand, A. Order of Play, Forward Induction, and Presentation Effects in Two-Person Games. *Experimental Economics* **2003**, *6*, 5–25.

Nagel, L. Unraveling in Guessing Games: An Experimental Study. *American Economic Review* **1995**, *85*(5), 1313–1326.

Neugebauer, T.; Sadreih, A.; Selten, R. Taming Selten's Horse with Impulse Response. *Working Paper*

Rapoport, A. Order of Play in Strategically Equivalent Games in Extensive Form. *International Journal of Game Theory* **1997**, *26*, 113–136.

Sadanand, A., Ideal Reactive Equilibrium. *Games* **2019**, *10*(2), 19.

Schotter, A.; Weigelt, K.; Wilson, C. A Laboratory Investigation of Multiperson Rationality and Presentation Effects. *Games and Economic Behavior* **1994**, *6*, 445–468.

Selten, R. Spieltheoretische Behandlung Eines Oligopolmodels Mit Nachfrageträgheit. *Zeitschrift für die gesampte Staatswissenschaft* **1965**, *121*, 301-324, 667-689.

Selten, R. Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games. *International Journal of Game Theory* **1975**, *4*(1), 25–55.

Stahl, D.O.; Wilson, P.W. Experimental evidence on players' models of other players. *Journal of Economic Behavior and Organization* **1994** *25*(3), 309–327.

Thompson, F.B. Equivalence of Games in Extensive Form. *Rand Research Memorandum* **1952**, *RM-759*, 1–13.

Appendices

A. Definition of Ideal Reactive Equilibrium

The definition of Ideal Reactive Equilibrium as presented in Sadanand (2019):

1. *Deliberate deviation of order* n, $n \geq 1$: Starting from any set of initial choices for all the players, a deliberate deviation of order n by a particular player at a particular information set is a departure by that player from his/her initial choice to another of his/her available choices (pure action, no randomization), and he/she supposes that n of his/her immediate followers can observe the deviation. That is, we erase the information sets (if there are any) for those n players' choices that follow the particular player.

   After a deliberate deviation, any player may respond with his/her best response to any changes that have occurred. This includes the original player. After that, any other player may respond with his/her best response to all the changes that have occurred thus far, and so on, until everyone is using a best response and we are at a (possibly different) Nash equilibrium.

2. *Successful deliberate deviation:* A deliberate deviation of any order is said to be successful if after the deviation and after all players are allowed to respond using best responses, the final outcome results in an increase in the payoff for the deviating player.

If a deliberate deviation is successful, then the outcome is that we remain at the final Nash equilibrium after deviation and best responses. If it is not successful, then the original deviator retracts his/her deviation, and we are back at the original starting point.

3. *Thought process:* A thought process is a sequence of outcomes starting with a deliberate deviation and then tracking all the best responses until we arrive at the resolution of that deviation. If the deviation is unsuccessful, then no movement occurs under the thought process; if it is successful, then the thought process tracks all the various stages of outcomes until we reach the final Nash equilibrium. Then, another deliberate deviation may be be undertaken by any player, and the thought process continues to track the outcomes of that deviation and all the best responses that follow, as before. This continues until either there are no more successful deviations available or until we observe that the thought process is cycling indefinitely through a subset of the terminal nodes.

4. **Ideal Reactive Equilibrium:** A Nash equilibrium is an ideal reactive equilibrium if:
   - all thought processes end with no change in the players' choices,
   - or if all thought processes cycle endlessly through the same subset of terminal nodes, then the original Nash equilibrium is IRE if either its outcome is part of the subset through which the thought process cycles or, in the case of the original Nash equilibrium being a mixed strategy Nash equilibrium, if the cycling utilizes all the actions with positive probability weights in the mixed strategy.

B. Equilibria of the alternative model

Nash equilibria of the alternative version of the original game:

- For $\alpha \in [0, 1)$ and $\beta \geq 1/4$: $p = 1$, $r \in [0, 1]$ and $q = 1$.
- For $\alpha \in [0, 1)$ and $\beta < 1/4$: $p = 1$, $r \leq \frac{2}{3-4\beta}$ and $q = 1$.
- For $\alpha < 1/2$ and $\beta = 1/4$: $p \in (1/3, 1)$, $r = 1$ and $q = 1$.
- For $\alpha + \beta < 1/12$: $p = 0$, $r = 0$ and $q \leq \frac{1}{4(1-\alpha-\beta)}$
- For $\alpha + \beta \in [1/12, 1/3]$: $p = 0$, $r = 0$ and $q \leq \frac{1-3(\alpha+\beta)}{3(1-\alpha-\beta)}$.
- For $\beta \leq 1/4$ and $\alpha + \beta < 3/4$ and $12(\alpha + \beta)(1 - \alpha - \beta) > 1 + 3\alpha - \beta$: $p = 1/3$, $r = 1$, $q = \frac{3(\alpha+\beta)}{1+3\alpha-\beta}$.

- For $\alpha+\beta \geq 1/12$ and $12(\alpha+\beta)(1-\alpha-\beta) \leq 1+3\alpha-\beta$: $p = \frac{(1-\alpha-\beta)[12(\alpha+\beta)-1]}{8\alpha-(1-\alpha-\beta)[12(\alpha+\beta)-1]}$, $r = \frac{(1-\alpha-\beta)[12(\alpha+\beta)-1]}{4\alpha}$ and $q = \frac{1}{4(1-\alpha-\beta)}$.

Nash equilibria of the alternative version of the variant:

- For $\alpha \in [0, 1)$ and $\beta = 0$: $p = 1$, $r = 2/3$ and $q = 1$ ($5ap_l \geq \max\{8\alpha - 6, 0\}$).
- For $\alpha \in [0, 1)$ and $\beta < \frac{2-\alpha}{3}$: $p = 1$, $r < 2/3$ and $q = 1$ ($p_l = 1$).
- For $\alpha \in [0, 1)$ and $\beta = \frac{2-\alpha}{3}$: $p = 1$, $r \leq \frac{2-3\beta}{3-3\beta}$ and $q \geq \frac{1-r(1-\beta)}{(3-4r)(1-\beta)}$ ($p_l = 1$).
- For $\alpha \in [0, 1]$ and $\beta > \alpha$: $p = 0$, $r = 0$ and $q = 0$ ($p_l = 1$).
- For $\alpha \in [0, 1]$ and $\beta = \alpha$: $p = 0$, $r = 0$ and $q \leq 1/4$ ($p_l = 1$).
- For $\alpha \in [0, 1)$ and $\beta = 1/2$: $p = 0$, $r < 2/3$ and $q = 1/4$ ($p_l = 1$).
- For $\alpha = 0$ and $\beta = 1/2$: $p = 0$, $r > 2/3$ and $q = 1/4$ ($p_l = 0$).
- For $\alpha = 0$ and $\beta = 1/2$: $p = 0$, $r = 2/3$ and $q = 1/4$ ($p_l \in [0, 1]$).
- For $\alpha \in (0, 3/5)$ and $\beta < \frac{3-5\alpha}{6}$: $p = 0$, $r < \frac{1-\alpha-2\beta}{\alpha}$ and $q = 1/4$ ($p_l \in [0, 1]$).
- For $\alpha \in (0, 3/5]$ and $\beta \in [\frac{3-5\alpha}{6}, \frac{1}{2}]$: $p = 0$, $r = 2/3$ and $q = 1/4$ ($p_l = \frac{5\alpha+6\beta-3}{5\alpha}$).
- For $\alpha \in (3/5, 1)$ and $\beta \leq \frac{1}{2}$: $p = 0$, $r = 2/3$ and $q = 1/4$ ($p_l = \frac{5\alpha+6\beta-3}{5\alpha}$).

C. Experimental Instructions (Translation from the original document in Greek).

The following instructions were those of Treatment 1. Instructions for Treatment 2 were totally analogous. We also provide the exact framing of the statements that were appearing to each player before making a choice, which might be important in an environment that employs the strategy method.

INSTRUCTIONS

Thank you for participating in this session. Please be quiet. The experiment will be conducted using a computer and all answers will be recorded by computer. Please do not talk to each other and keep quiet during the session. The use of mobile phones and other electronic devices is not permitted. Please carefully read the instructions and if you have any questions, raise your hand and the answer that will be given so as to be heard by everyone.

General Instructions

The experiment includes a series of choices to be made. Each of you will receive a financial reward. **The exact amount you will receive will depend on both the decisions you will make during the experiment and the decisions of other participants.** Additional information regarding the payment can be found below. In addition, you will receive the amount of 5 euros as a financial reward for participating. Following the completion of the experimental session, you will be paid privately in cash, by giving us the id-number of the computer that you used.

There is no time limit for submitting your choices. However, it is recommended that the choices should be made within a reasonable time frame. If asked by the researchers, please submit your choices as soon as possible.

The Experiment

The experiment will consist of 50 periods. Each period will be completely independent of the others. In each period, participants will be split into groups of three and the members of each group will interact only among themselves. The groups will change in each period, hence the people you will interact in each period will most probably be different than those you interacted in the previous period.

The members of each group will be playing a simple game between them. More specifically, the three players will choose **sequentially** between two available moves. The profit of each player will be counted in points and **will depend both on his own choice and on the choice of the other two players.**

The order of moves is the following:

1. The **FIRST** player choose one of the moves **A or B**:
   - If the **FIRST** player chooses **A**, then the choice of the **SECOND** player **WILL NOT AFFECT** the result of the game.
2. Next, the **SECOND** player will choose one of the moves **A or B**, *for the case in which the* **FIRST** *player chose* **B**.
   - If both the **FIRST** and the **SECOND** player choose move **B**, then the choice of the **THIRD** player **WILL NOT AFFECT** the result of the game.

3. Finally, the **THIRD** player chooses one of the moves **A or B**, *for the case in which either the* **FIRST** *or the* **SECOND** *player have chosen move* **A***, without knowing which of the two has happened.*

4. According to the choices of the three players the profits of each player from the given round are realized. In the following pages there is a figure that describes the sequence of moves and the profits of each player for each combination of choices, as well as some explanatory examples.

The role of each participant will be determined randomly between the members of the group of three, in the beginning of the period. Therefore, all participants will be require to play all three roles (**FIRST, SECOND and THIRD player**) in some of the periods.

Payment

At the end of the experiment, the points you collected during the 50 periods will be calculated and their value in euros will be equal to the number of points divided by 100. To those will be added the 5 euros of the participation fee.

$$\text{Payment } = 5 + \text{ (Points you collected) } /100$$

1. If the **FIRST** player chooses move A, then the **SECOND** player chooses either move A or move B and finally the **THIRD** player chooses move A, then the profits of the three players will be:
   (a) *30 points for the***FIRST** *player*
   (b) 30 points for the **SECOND** *player and*
   (c) 20 points for the **THIRD** *player.*

2. If the **FIRST** player chooses move A, then the **SECOND** player chooses either move A or move B and finally the **THIRD** player chooses move B, then the profits of the three players will be:
   (a) *0 points for the* **FIRST** *player*
   (b) 0 points for the **SECOND** *player and*
   (c) 0 points for the **THIRD** *player.*

3. If the **FIRST** player chooses move B, then the **SECOND** player chooses move A and finally the **THIRD** player chooses move A, then the profits of the three players will be:
   (a) *40 points for the* **FIRST** *player*
   (b) 40 points for the **SECOND** *player and*
   (c) 0 points for the **THIRD** *player.*

4. If the **FIRST** player chooses move B, then the **SECOND** player chooses move A and finally the **THIRD** player chooses move B, then the profits of the three players will be:

   (a) *0 points for the* **FIRST** *player*

   (b) 0 points for the **SECOND** *player and*

   (c) 10 points for the **THIRD** *player.*

5. If the **FIRST** player chooses move B, then the **SECOND** player chooses move B and finally the **THIRD** player chooses either move A or move B, then the profits of the three players will be:

   (a) *10 points for the* **FIRST** *player*

   (b) 10 points for the **SECOND** *player and*

   (c) 10 points for the **THIRD** *player.*

End of instructions.

FRAMING OF QUESTIONS (Translation from the original text in Greek.)

- *(Treatment 1, FIRST Player)*:

  In this round you are the **FIRST** player.

  Which one of the following moves do you choose? [Buttons] **A** [or] **B**.

- *(Treatment 1, SECOND Player)*:

  In this round you are the **SECOND** player.

  The **FIRST** player has already played.

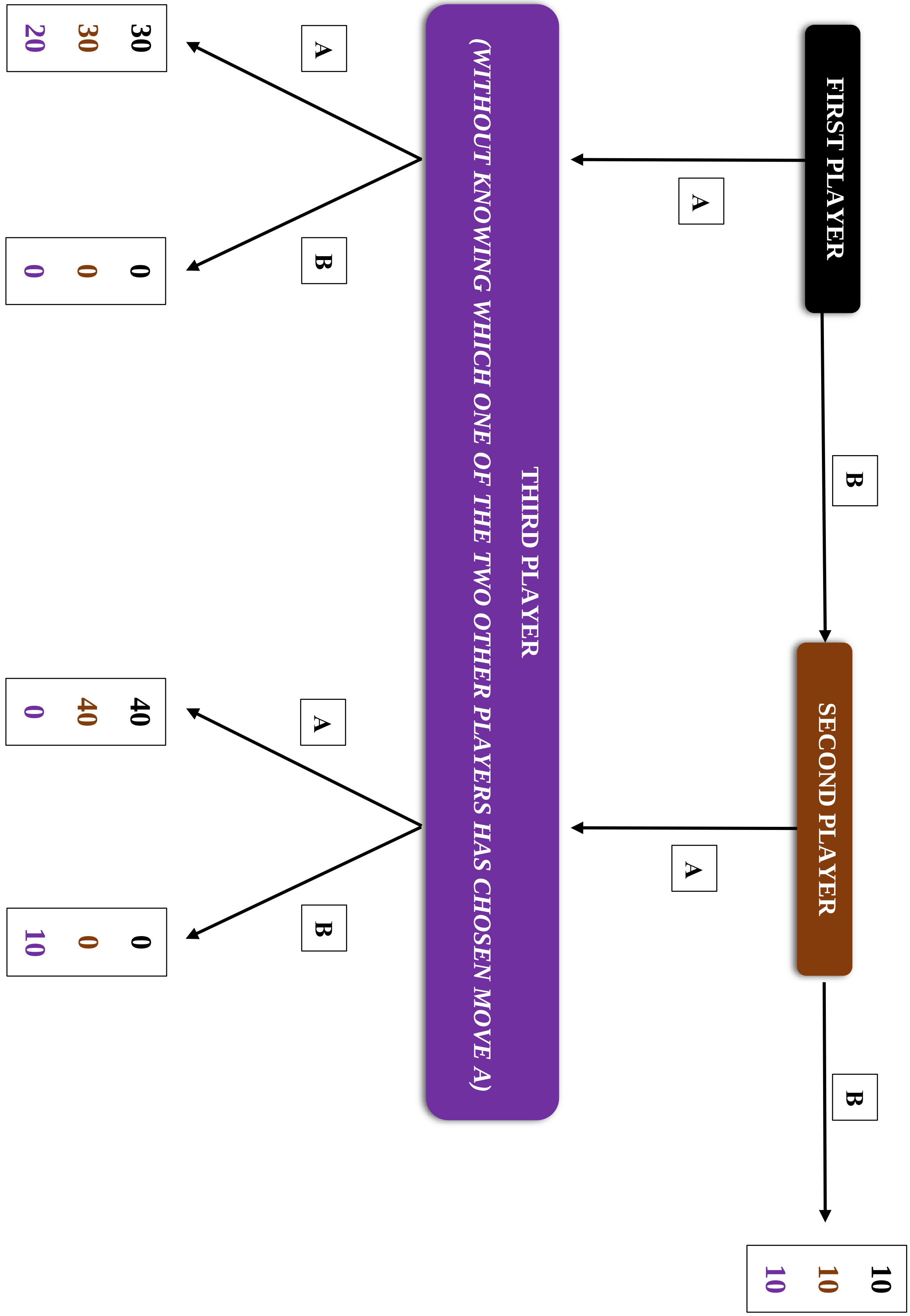
  Your move will affect payoffs ONLY if the **FIRST** player has chosen move **B**.

  Which one of the following moves do you choose, if the **FIRST** player chose move **B**? [Buttons] **A** [or] **B**.

- *(Treatment 1, THIRD Player)*:

  In this round you are the **THIRD** player.

  The **FIRST** and the **SECOND** player have already chosen.



FIRST PLAYER
A
B
SECOND PLAYER
A
B
10 10 10
THIRD PLAYER
(WITHOUT KNOWING WHICH ONE OF THE TWO OTHER PLAYERS HAS CHOSEN MOVE A)
A
B
A
B
30 30 20
0 0 0
40 40 0
0 0 10

Which one of the following moves do you choose, if at least one of the first two players chose move **A**? [Buttons] **A** [or] **B**.

- *(Treatment 2, FIRST Player – respective to THIRD player of Treatment 1)*:

  In this round you are the **FIRST** player.

  Which one of the following moves do you choose? [Buttons] **A** [or] **B**.

- *(Treatment 2, SECOND Player – respective to FIRST player of Treatment 1)*:

  In this round you are the **SECOND** player.

  The **FIRST** player has already played.

  Which one of the following moves do you choose, if the **FIRST** player chose move **B**? [Buttons] **A** [or] **B**.

- *(Treatment 2, THIRD Player – respective to SECOND player of Treatment 1)*:

  In this round you are the **THIRD** player.

  The **FIRST** and the **SECOND** player have already chosen.

  Which one of the following moves do you choose, if the **SECOND** player chose action **B**? [Buttons] **A** [or] **B**.

Part I

title