Naivete and Sophistication in Initial and Repeated Play in Games*

Bernardo García-Pola† Nagore Iriberri†‡

September 17, 2020

Abstract

Compared to more sophisticated equilibrium theory, naive, non-equilibrium behavioral rules often better describe individuals' initial play in games. Additionally, in repeated play in games, when individuals have the opportunity to learn about their opponents' past behavior, learning models of different sophistication levels are successful in explaining how individuals modify their behavior in response to the provided information. How do subjects following different behavioral rules in initial play modify their behavior after learning about past behavior? This study links both initial and repeated play in games by analyzing elicited behavior in 3×3 normal-form games using a within-subject laboratory design. We classify individuals into different behavioral rules in both initial and repeated play and test whether and/or how strategic naivete and sophistication in initial play correlate with naivete and sophistication in repeated play. We find no evidence of a positive correlation between naivete and sophistication in initial and repeated play.

Keywords: Naivete, sophistication, strategic thinking, initial play, repeated play, level-k thinking, adaptive and sophisticated learning, mixture-of-types estimation

*This project was supported by a 2018 Leonardo Grant for Researchers and Cultural Creators, BBVA Foundation. The Foundation accepts no responsibility for the opinions, statements and contents included in the project and/or the results thereof, which are entirely the responsibility of the authors. We would also like to thank Vincent P. Crawford, Pedro Rey-Biel and seminar attendees of various presentations for their useful comments.

†AGORA Center, Dept. of Economics, UNSW Business School, Sydney, NSW 2052, Australia (bernardo.garciapola@gmail.com).

‡University of the Basque Country, EHU-UPV, and IKERBASQUE, Basque Foundation for Research (nagore.iriberri@gmail.com)

1 Introduction

Nash equilibrium has been and still is the benchmark solution concept in game theory for predicting individual behavior in strategic environments. Since economics adopted the use of laboratory experiments, hundreds of experimental studies have tested whether individual behavior complies with the predictions of Nash equilibrium theory. These studies have shown that equilibrium theory has clear limitations in regard to its ability to describe how people behave in strategic environments. In response to ample experimental evidence, the important contributions of behavioral economics include models of bounded rationality that improve our understanding of how people actually behave in two different domains. First, when individuals make decisions for the first time with no previous experience or opportunity to learn, a scenario that is called initial play, naive, non-equilibrium, behavioral rules often exceed equilibrium theory in their ability to describe individual behavior (see for example, Goeree and Holt, 2001, and Crawford et al., 2013).[1] Second, given that people often do not start playing the Nash equilibrium strategy, bounded rationality models have been applied to repeated play to understand how people modify their behavior when provided with information on past behavior, that is, models that explain how individuals learn over time (see for example, Sobel, 2000).

Does behavior in initial responses relate in any way to behavior in repeated play in strategic environments? This is the central question of this paper.

When studying initial play, models that explain how individuals begin playing games differ in the naivete or sophistication assumed with respect to individual thinking in strategic environments. We can order the behavioral rules in initial play from most naive to most sophisticated.[2] We propose that the most naive behavioral rules include processes that require no *strategic* thinking, meaning *no need* to predict oppo-

[1]For initial play, the crucial aspect is that individuals make decisions with no previous experience or opportunity to learn. This can include responses to one-shot games, as in Goeree and Holt (2001), or responses to multiple games that are similar but in which subjects are not provided with any feedback from game to game, as in, for example, Costa-Gomes et al. (2001).

[2]We use naivete and sophistication to refer to *strategic* naivete and sophistication, which can be different from *behavioral* naivete and sophistication. In other words, an individual showing behavior consistent with level-*1* behavioral rule is naive regarding the revealed *strategic* sophistication but can be *behaviorally* the most sophisticated if all other opponents show random uniform behavior. This observation is related to work by Alaoui and Penta (2016) that tests whether individuals who show behavior consistent with a particular level-k are due to their own limitations or to their beliefs about opponents' limited behavior.

nent's behavior, such that strategic settings are considered to be isomorphic to pure decision making settings. For example, maxmax (optimistic) and maxmin (pessimistic) behavioral rules fall into this category because maximizing over possible outcomes or maximizing over minimum possible outcomes does not require any ability to predict an opponent's behavior. Level-k thinking models, which have been shown to be successful in explaining initial behavior in different settings (Stahl and Wilson, 94, 95; Nagel, 95; Costa-Gomes et al., 2001, Camerer et al., 2004), illustrate different levels of strategic sophistication quite well. Level-*1* behavioral type calculates the expected payoff associated with each of the available strategies, assuming that each of the opponent's actions is equally likely, and takes the strategy with the highest expected payoff; alternatively, sums own payoffs across columns and takes the strategy that yields the maximum expected payoff. In the spirit of this latter interpretation, we also consider level-*1* to be a naive behavioral model.[3] More sophisticated behavioral rules require individuals to best respond to some type of opponent behavior. Level-*2* and level-*3* represent assumptions of increasing sophistication about the opponent's actions, as level-*2* believes the opponent is behaving as a naive level-*1* and best responds to those beliefs, while level-*3* assumes the opponent behaves as a level-*2* and best responds to those beliefs. Finally, among the most sophisticated behavioral rules is the Nash equilibrium, which considers not only common knowledge of rationality but also rational expectations about beliefs.

In studies focused on repeated play, models that explain how individuals modify their behavior in response to information on (own and opponent's) past behavior also differ in the strategic naivete or sophistication with respect to whether individuals use information on past behavior and, if they do how they use it. Learning models can also be ordered according to their sophistication level from most naive to most sophisticated in a hierarchical manner. We propose that the most naive learning model is the one that simply repeats the same strategy used in the past, having *no need* to use opponent's past strategy. We refer to this as the *No-Change* behavioral rule in repeated play. Adaptive learning models assume that individuals modify their behavior in response to information on past behavior, i.e., best responding to an opponent's past behavior

[3] The cognitive hierarchy model (Camerer et al., 2004) assumes that level-k players best respond to combinations of existing lower levels. However, both level-k thinking and cognitive hierarchy models coincide in terms of the level-*1* predictions.

(illustrated best by fictitious play, as in Fudenberg and Levine, 98a and 98b). Note that adaptive learners assume that opponents indeed follow a *No-Change* type, as they assume that opponents will repeat the same strategy used in the past; therefore, adaptive learners will best respond to their opponents' past strategy. Finally, more sophisticated learning models assume that opponents indeed learn through an adaptive learning model and accordingly best respond to this (see, for example, Milgrom and Roberts, 91; Selten, 91; Conslik, 93a and 93b; Nagel, 95; Camerer et al., 2002, and Stahl, 2003).

Somewhat surprisingly, the literature on learning models (i.e., Cheung and Friedman, 97; Erev and Roth, 98, Fudenberg and Levine, 98a and 98b, and Camerer and Ho, 99) and the literature on models to explain initial behavior (summarized in Crawford et al., 2013) have followed parallel paths.[4] On the one hand, when studying learning over time, initial play has been treated as a "black box", an exogenous factor used only to initialize learning models, such as by estimating initial "attractions" associated with each of the particular strategies or, alternatively, simply assuming that initial "attractions" are the same across different strategies. On the other hand, models that aim to explain initial behavior have used mostly experimental designs that provide no feedback from game to game, precisely to suppress any opportunity to learn. Such models have been silent on explaining learning over time.

However, it appears to be natural that some type of relation exists between strategic behavior in initial and repeated play. As Costa-Gomes and Crawford (2006) note, modeling initial responses more precisely could yield insights into cognition that elucidate other forms of strategic behavior, such as learning and distinguishing between different levels of sophistication in rules and therefore influencing implications for equilibrium selection and convergence. However, similar consistencies that seemed a priori intuitive have been empirically rejected (Costa-Gomes and Weizsacker, 2008; Knoepfle et al., 2009). Gill and Prowse (2016) take a different approach by measuring cognitive ability exogenously with the Raven Test; they then test if more cognitively able subjects choose numbers closer to equilibrium, converge more frequently to equilibrium play and earn more. The question of whether the behavior in these two contexts is related

[4]There are a few exceptions, as some models have been used to explain both initial behavior and learning behavior over time, such as quantal response equilibrium by McKelvey and Palfrey (95), which simply estimates different noise levels or lambda-s for behavior in different stages.

arises not only as a natural question but also as an important one. If such behavior is related, observing the initial behavior of an individual would be informative of how her behavior will change and vice versa. Furthermore, this relation would allow a unified framework of behavior in games that incorporates both initial and repeated play (see, for example, Ho et al., forthcoming). If such behavior is not related, such that we observe very different levels of sophistication when the same individual faces a situation for the first time and in repeated play, characteristics that we sometimes measure as inherent to an individual, such as cognitive ability, may be more context-dependent than previously believed.

We therefore study fundamental questions for proposing a unified framework for studying initial and repeated play in games: How do strategic naivete and sophistication in initial play relate to naivete and sophistication in the use of information on past behavior in repeated play? Is a strategically naive player in initial responses, compared with a more sophisticated player, more likely to learn through a naive learning model in repeated play? We propose a laboratory experiment and a mixture-of-types model econometric estimation to address these inherently empirical questions.[5]

The subjects in our experiment proceeded through 14 different 3×3 games (actually 7 asymmetric games, where the subjects play both as row and column players) two times in two different stages of the experiment. In the first stage, the subjects receive no feedback from game to game, with the objective of eliciting their initial play (with no opportunity to learn or obtain experience from game to game). Subjects' behavior could not have been affected by anything in the second stage, as they did not know what they would do in the second stage. Based on the subjects' profiles of 14 decisions, we classify each subject as following one of multiple behavioral rules. This exercise is similar to those pioneered by Stahl and Wilson (94, 95) and later used by, for example, Costa-Gomes et al. (2001), Costa-Gomes and Crawford (2006), Rey-Biel (2009) and García-Pola et al. (2020).

In the second stage of the experiment, the subjects repeat the same 14 games, but this time, in each of the games, they receive information on what they did in the first

[5]As initial response can be equivalent to response to one-shot games, our study can also be phrased as testing if naivete/sophistication in behavior in one-shot games correlates in any way with naivete/sophistication in behavior in one-shot games when provided with information from past behavior. This is related to our design being able to capture "initial model of learning", as we refer to the objective of stage 2 of the experiment we carry out.

stage, as well as what their *current* opponent did in the first stage. Using the subjects' profiles of 14 decisions and observed information on their own and current opponent's past strategies, we classify each subject as following one of multiple behavioral rules in repeated play.

It is important to note that this elicitation and identification of learning rules is different from studies that attempt to identify the ability of different learning rules to explain behavioral data (see, for example, Erev and Roth, 98; Camerer and Ho, 99; Feltovich, 2000, and more recently, Kovarik et al., 2018). First, the learning models we consider and identify differ among them on which information (own or opponent's) individuals use for modifying past behavior and on what individuals believe about how the opponents will use that same information on past behavior. Second, in our setting, for a particular game, subjects can learn about an opponent's past actions just once, but we elicit how subjects learn from 14 different games or decisions based on their opponents' past actions in those 14 different games. In other words, we elicit subjects' learning rules using multiple different games in a way that does not allow the subjects themselves to evaluate how successful their learning model is, which we refer to as the "initial model of learning". These two important features considerably distinguish our approach to studying learning from existing work.

As this study is, as far as we know, the first empirical exercise to connect initial and repeated play, we designed games with the purpose of allowing for the highest separation among different behavioral rules in both initial and repeated play. The separation is the cornerstone for the use of a mixture-of-types model when identifying and classifying subjects into different behavioral rules both in initial and repeated play. Finally, the within-subject design allows us to construct contingency tables to test whether naivete and sophistication in initial play are correlated with naivete and sophistication in repeated play.

We find no evidence for a positive correlation between strategic naivete and sophistication in initial and repeated play. Regarding initial behavior, consistent with previous findings, we find that the majority of subjects, 60%, use a naive, non-strategic, behavioral rule. The second most frequent rule is a more sophisticated behavioral rule, level-*2*, used by 36% of the subjects. Additionally, consistent with previous findings, few Nash equilibrium players are found among the subject population. Furthermore, when identifying the behavioral rules that describe repeated play, the majority of subjects, approximately 50%, show behavior that is consistent with adaptive learning and an important number of subjects, approximately 40% of them, follow the most naive behavioral rule of ignoring their opponent's past action. Sophisticated learning models are also rarely used. Most importantly and surprisingly, when naivete and sophistication are compared between initial and repeated play, which is the central question of our study, we find little support for any positive correlation. Subjects using a naive behavioral rule in initial play are if anything more likely to use a more sophisticated learning model than subjects using a more sophisticated model in initial play. In particular, 60% of individuals using a naive behavioral rule in initial play use an adaptive learning model, while only 40% of the level-*2* use an adaptive learning model.

The rest of this paper is organized as follows. Section 2 describes the experimental procedures and design in detail. Section 3 presents the theoretical framework, describing behavioral rules and their predictions. We also define naivete and sophistication in behavioral rules, assessing the experimental design. Section 4 presents the results, which are divided into the identification and classification of subjects according to their initial play, identification and classification of subjects according to their repeated play, and the correlation between naivete and sophistication across the two settings. Section 5 includes three important robustness checks of the potential misspecification in the identification and classification of behavioral rules. Finally, Section 6 concludes the paper.

2 Experimental Procedures and Design

2.1 Procedures

A total of 198 subjects who participated in the experiment were recruited using the ORSEE system (Greiner, 2015). The sessions were conducted via computer using z-Tree software (Fischbacher, 2007). In April and May 2019, two sessions with a total of 78 subjects took place in the Laboratory of Experimental Analysis (Bilbao Labean; http://www.bilbaolabean.com) at the University of the Basque Country. We conducted two additional sessions with the remaining 120 subjects in the Laboratory of Experimental Economics (LEE, http://lee.uji.es) at the University Jaume I of Castellón.

The subjects were told that the experiment consisted of two different parts and that payments would depend on both luck and their own and other subjects' decisions. Immediately before each part, the subjects were given detailed instructions explaining the task involved, including examples of games, how they could make decisions, and how they were going to be matched and paid. The subjects were allowed to ask any questions they might have during the instructions. At the end of the instructions, the subjects were asked a few questions to guarantee that they understood the instructions regarding each part. They could not start the experiment until they answered those questions correctly. A translated version of the instructions can be found in Appendix C.

All subjects played the same seven 3×3 normal-form two-player games in the same order, first as the row player and then as the column player, playing a total of 14 games in each part.[6] We did not inform the subjects that they played the same games in different roles, and we showed all the games to all subjects from the perspective of row players. The subjects were randomly matched in a way that, within each part of the experiment, they were paired with a different opponent in each of the 14 games. In the first part of the experiment, the subjects received no feedback from game to game to elicit initial play in the 14 games. In the second part, the subjects repeated the same 14 games in the same order but were provided with information about their own past strategy and their current opponent's past strategy in the first part of the experiment. The fact that subjects were provided with information on past actions in the second part was public knowledge, but they *only* learned about the availability of information after they had finished the first part. In other words, the behavior in the first part of the experiment could not be affected in any way by any experimental feature in the second part. An example of how the games in both parts and the provided information in the second part were displayed in the experiment can be found in the instructions in Appendix C.

When all subjects had submitted their choices in the two parts, for each subject, the computer randomly chose two games from any of the two parts for payment. Thus, each

[6] Any experiment using a within-subject design may worry about potential experimenter demand effects. However, it is not clear to us how experimenter demand effects may have affected the results in our setting. On the one hand, individuals' natural taste for consistency in behavior (Eyster, 2003; Falk and Zimmermann, 2011) may lead subjects to show the same behavior in both parts of the experiment. On the other hand, if the subjects had identified our research question, one might anticipate more reaction to the provided information. The results clearly show that not all subjects repeated the same behavior and that not all subjects reacted to the provided information.

subject could be paid for different games. Before being paid, the subjects completed a non-incentivized questionnaire regarding demographic data (gender, age, nationality, university entry grade and field of study), risk preferences following Eckel and Grossman (2002), a cognitive reflection test. Descriptive statistics of all these variables can be found in Appendix Table A1. The subject pool showed the typical characteristics of undergraduate students who are mostly studying for economics and business degrees, with a slightly higher presence of females, given that most were pursuing a degree in social sciences. We also requested free-format explanations of their choices and the expected choices of others in each of the parts of the experiment. We did not include these data in the analysis, but we did informally assess the consistency between the subjects' explanations of what they did and the rule we estimated using their elicited actions and frequently observed a clear coincidence between the two. For a work that attempts to relate subjects' free-format explanations of their actual actions and their actions, see Brañas et al. (2011). Finally, we paid the subjects privately according to the two games selected plus a 3 Euro show-up fee. The average payment was 15.76 Euros, with a standard deviation of 4.90. The entire experiment lasted one hour and a half, including the reading of instructions and payment.

2.2 Design of Games

We designed seven 3×3 normal-form games, as shown in Figure 1. The actual order in which the games were presented to the subjects was G1, G2... until G7 as row players, to which we will refer as G11, G21, and so on until G71, and G1, G2... until G7 as column players, to which we will refer as G12, G22, and so on until G72. As noted in the previous section, all subjects were presented the games as if they were row players, that is, we transposed the games when the subjects were playing as column players. We chose this particular sequence, first as row players and then as column players, to prevent the subjects from realizing that they were making choices in the same games.

We chose 3×3 normal-form games because such games allow for ample separation between the predictions of different behavioral rules. Note that with 14 3×3 games, there are 4,782,969 possible ways of playing the 14 games, while with 2×2 games, we would have *only* 16,384 possible combinations. Therefore, having 3×3 games substantially increases the a priori possibility of separation among the predictions of

Figure 1: Experimental Games

G1

4 , 20	20 , 12	18 , 2
6 , 8	8 , 14	22 , 16
18 , 14	14 , 6	2 , 18

G2

6 , 18	22 , 4	4 , 16
20 , 6	2 , 24	16 , 4
12 , 12	2 , 6	18 , 22

G3

4 , 20	12 , 16	16 , 4
18 , 2	20 , 12	2 , 8
22 , 18	2 , 2	10 , 22

G4

10 , 18	20 , 16	4 , 6
12 , 10	14 , 22	2 , 12
6 , 4	18 , 4	16 , 18

G5

8 , 16	16 , 14	20 , 12
16 , 8	18 , 12	4 , 4
14 , 6	16 , 4	2 , 20

G6

14 , 16	2 , 20	12 , 22
6 , 18	20 , 4	10 , 6
22 , 4	14 , 18	4 , 10

G7

4 , 20	22 , 14	18 , 4
6 , 6	8 , 12	20 , 14
18 , 16	14 , 8	4 , 18

different behavioral rules. Additionally, we chose 3×3 games instead of, for example, 4×4 games to ensure that the number of strategies was relatively small such that it was easier to handle by subjects, which facilitated the explanation of the instructions. Finally, we designed our own games instead of using games from other studies because we aimed to have high separation between different behavioral rules both in initial and repeated play, which, as far as we know, was not the aim of any previous study.

3 Theoretical Framework

We now explain in detail which behavioral rule we consider in each of the two parts of the experiment and describe the predicted strategies across the 14 games. We also define how we classify the behavioral rules into the spectrum of naivete and sophistication. We finish this section by showing the actual separation between the predicted behavior.

3.1 Naivete and Sophistication in Initial Play and Repeated Play: Predictions

When analyzing initial play, we consider 8 behavioral types. We take the leading behavioral models in the literature (Stahl and Wilson, 94 and 95, Nagel, 95, Costa-Gomes et al. 2001, Costa-Gomes and Crawford, 2006, and García-Pola et al., 2020, among others).

The altruistic or social welfare maximizer type, *A*, simply sums own and opponent's payoffs in each cell of the payoff matrix and applies the maxmax operator. The inequity averse type, *IA*, in a similar way, takes the absolute value of the difference between the own and opponent's payoffs in each cell of the payoff matrix and applies the minmin operator. Although these two models resonate with interdependent preferences, which a priori are independent from models of strategic thinking, the actual naive implementation brings them close to a naive behavioral rule. The optimistic type (*MaxMax*) follows the strategy that results from applying the maxmax operator using only own payoffs, while the pessimistic type (*MaxMin*) follows the strategy that results from applying the maxmin operator using only own payoffs. The level-*1* type (*L1*) sums own payoffs across columns (opponent's three possible strategies) and takes the strategy that yields the maximum sum of payoffs. Level-*2*, *L2*, expects the opponent to behave as level-*1* type and best respond to those beliefs. Level-*3*, *L3*, similarly, expects the opponent to behave as level-*2* type and best responds to those beliefs.[7] Finally, *NE* play calculates the mutual best response required by equilibrium thinking. The top panel of Table 1 shows the predictions of each of these behavioral rules in the 3×3 normal-form games in Figure 1.

[7] Note that given that our games in Figure 1 do not have any dominated strategies, level-k rules and dominance-k rules, as defined in Costa-Gomes et al. 2001, coincide.

Table 1: Predicted Strategies by Different Behavioral Rules

	G11	*G12*	*G21*	*G22*	*G31*	*G32*	*G41*	*G42*	*G51*	*G52*	*G61*	*G62*	*G71*	*G72*
Initial Play														
A	2	3	3	3	3	1	1,2	2	1	3	1	3	1	2
IA	2	1	3	1	3	2	1,2,3	1,3	2	3	1	1	2	1
MaxMax	2	1	1	2	3	3	1	2	1	3	3	3	1	1
MaxMin	2	1	1	1	1	3	3	3	1	1	2	3	2	2
L1	1	1	2	3	2	1	3	2	1	3	3	2	1	1
L2	3	1	3	2	3	2	1	3	1	1	2	2	3	1
L3	3	3	1	3	2	3	3	1	2	1	2	1	3	3
NE	2	3	3	3	2	2	3	3	2	2	1	3	2	3
Repeated Play														
No-Change	"Same strategy as in the first part"													
$Adaptive_S$	"Best response to (opponent's past strategy)"													
$Adaptive_A$	"*A* best response to (opponent's past strategy)"													
$Adaptive_{IA}$	"*IA* best response to (opponent's past strategy)"													
Sophisticated	"Best response to (opponent's best response to (own past strategy))"													
Sophisticated 2	"Best response to (opponent's best response to (best response to (opponent's past strategy)))"													

Notes: The table reports the strategies predicted by the models in initial play (top panel) and the models in repeated play (bottom panel); 1, 2 and 3 refer to the first, second and third strategies, respectively. In a few instances, a behavioral rule is indifferent between multiple strategies, so we assume the behavioral rule will predict any of those strategies with equal probability.

How should we classify all these behavioral rules from most naive to the most sophisticated? We take a simple approach and define the most naive behavioral rules as those that do *not need* to anticipate the opponent's strategy, such that subjects following the most naive behavioral rule could treat strategic and pure decision making situations as isomorphic. Among the non-strategic behavioral rules, we consider altruistic or social welfare maximizer, inequity-averse, maxmax or optimistic, maxmin or pessimistic and level-*1* type. Take into account that some of these behavioral rules can indeed be interpreted as individuals having beliefs and best responding to them (i.e., level-*1* or maxmin rule), but they can also be interpreted by simply following a naive behavioral rule as if individuals faced a pure decision making setting (i.e. level-*1* summing own payoffs across columns or maxmin doing maxmin operator over own payoffs). As long as a behavioral rule does not need to anticipate the opponent's strategy, we consider these behavioral rules to be non-strategic and naive. Once we define the most naive behavioral rule, we build on best response iterations to define higher levels of sophistication. In this way, level-*2* and level-*3* are ordered immediately after the naive behavioral rules because level-*2* anticipates the opponent will behave as a level-*1* and best responds to those beliefs, while level-*3* anticipates the opponent will behave as a level-*2* and best responds to those beliefs. Please see robustness test in Section 5.1 to see similar hierarchical best response iterations for *A*, *IA* and *MaxMax* and *MaxMin*. Finally, the most sophisticated behavioral rule is the Nash equilibrium.

In repeated play, we consider 4 main behavioral types. We also take the leading behavioral models from the literature (Fudenberg and Levine, 98a and 98b; Nagel, 95; Camerer et al., 02; Stahl, 03).

The no-change type (*No-Change*) simply mimics the behavior taken in the first part of the experiment. Adaptive learning behavior (*Adaptive*) assumes that individuals best respond and that they try to guess what their opponent will do (similar to any belief-based learning model, as in Fudenberg and Levine, 1998). In our setting, as subjects are provided with the opponent's past strategy, adaptive learning assumes that the opponent will repeat her/his past strategy, so opponents are expected to follow a *No-Change* type and therefore adaptive learners best respond to such behavior.[8] When best responding to opponents' past behavior, individuals can be maximizing their own payoffs ($Adaptive_S$), doing maxmax over the sum of their own and opponents' payoffs ($Adaptive_A$), and doing minmin over the absolute difference between their own and opponents' payoffs ($Adaptive_{IA}$). Sophisticated learning (*Sophisticated*) goes one step further and considers that the opponent follows adaptive learning behavior. As such, the sophisticated learning rule uses own past behavior, calculates the corresponding adaptive learning behavior (i.e., best response to own past behavior), and then best responds to those beliefs regarding the opponent's expected behavior. Finally, we also consider one more round of sophistication in repeated behavior (*Sophisticated 2*). The *Sophisticated 2* learning type assumes that the opponent follows sophisticated learning

[8] Notice that in our repeated play setting, given that subjects are never provided with how successful their past strategy in the first stage was, reinforcement learning (Erev and Roth, 98) cannot be directly assessed. However, with a more flexible interpretation and assuming that subjects evaluate their past strategy with the current opponent's past strategy, reinforcement and adaptive learning models would predict the same strategy.

behavior (best response to own behavior as an adaptive learner) and best responds to those beliefs. Note that all these behavioral types not only require a particular game to make predictions but also need own and/or opponents' past behavior, so they are dependent on observed past behavior. The bottom panel of Table 1, therefore, does not show the actual predicted strategies, but in general, the calculation of a particular behavioral rule requires repeated play with the provided information.

How should we classify all these behavioral rules from most naive to the most sophisticated? We again take a simple approach and define the most naive behavioral rule in a repeated play setting as the one does *not need* to use information on past strategies. Among the learning rules we consider, the *No-Change* type is therefore the most naive, that is, the one that simply repeats the strategy taken in the first stage. The rest of the behavioral rules build on this basis, increasing in sophistication as they take one additional step of best response iteration on the use of information, such that *Adaptive* rule is more sophisticated than the *No-Change* rule because adaptive learners are best responding to *No-Change* types, and *Sophisticated* rule is more sophisticated than the *Adaptive* rule because sophisticated learners are best responding to adaptive learners. Finally, the most sophisticated learning rule, *Sophisticated 2*, assumes that the opponent is a sophisticated learner and best responds to those beliefs.

3.2 Assessing the Design: Behavioral Rules' Separation across Games

The 3×3 normal-form games, shown in Figure 1, were carefully designed with the aim of having the largest separation between the predictions of different behavioral rules.

Table 2 shows the separation between different behavioral rules in both initial play (panel A) and repeated play with information on past actions (panel B). The values in the table represent the proportion of games in which the predictions of two behavioral rules (the one in the row and the one in the column) are separated. The numbers can take any value between 0 (no separation at all, such that two behavioral rules predict exactly the same strategy in each of the 14 games) and 1 (full separation, such that two behavioral rules predict a different strategy in each of the 14 games).

Table 2: Separation of Different Behavioral Rules

Panel A: Initial Play

	A	*IA*	*MaxMax*	*MaxMin*	*L1*	*L2*	*L3*
A	0.00						
IA	0.60	0.00					
MaxMax	0.46	0.62	0.00				
MaxMin	0.71	0.65	0.57	0.00			
L1	0.57	0.76	0.50	0.79	0.00		
L2	0.75	0.58	0.57	0.64	0.71	0.00	
L3	0.86	0.80	0.86	0.64	0.79	0.71	0.00
NE	0.57	0.51	0.86	0.64	0.79	0.79	0.57

Panel B: Repeated Play

	No Change	$Adaptive_S$	$Adaptive_A$	$Adaptive_{IA}$	*Sophisticated*
No Change	0.00				
$Adaptive_S$	0.65	0.00			
$Adaptive_A$	0.60	0.52	0.00		
$Adaptive_{IA}$	0.62	0.81	0.64	0.00	
Sophisticated	0.71	0.60	0.71	0.62	0.00
Sophisticated 2	0.71	0.60	0.50	0.70	0.47

Notes: The table reports the proportions of strategies across all 14 games in which the different behavioral models predict different strategies. The minimum possible separation value is 0, which occurs when the two models prescribe the same strategy in all 14 games, and the maximum possible separation value is 1, which occurs when the two models predict a different strategy in each of the 14 games.

The separation values for the initial play range between 0.46 and 0.86, which shows that each pair of behavioral rules is separated in at least 6 of 14 games and as many as 12 of 14 games. Regarding the separation values in repeated play, we could not calculate these values ex ante, as they depended on the particular observed past behavior of subjects.[9] The values in panel B are therefore based on the actual observed behavior in the first part of the experiment. The values range between 0.47 and 0.81, which indicates that two behavioral rules for repeated play are separated in at least 6 of 14

[9]We could indeed use, as we actually did, the accumulated evidence from past studies (see Crawford et al., 2013, for example) that found that approximately half of the subject population showed non-strategic behavior and a smaller proportion more sophisticated behavioral rules such as *L2* and *L3*, with a minority of subjects following the Nash equilibrium strategy. As the results in Section 4.2 show, we find a type distribution that is roughly consistent with existing findings in the literature.

games and as many as almost 12 of 14 games. We therefore conclude that the goal of attaining a large separation between the considered behavioral rules was achieved.

4 Results

4.1 Descriptive Overview

We begin by considering the mean behavior in both initial and repeated play, which represents how individuals start playing in strategic environments with no feedback (first part) and how individuals react to both their own and current opponent's past behavior (second part).

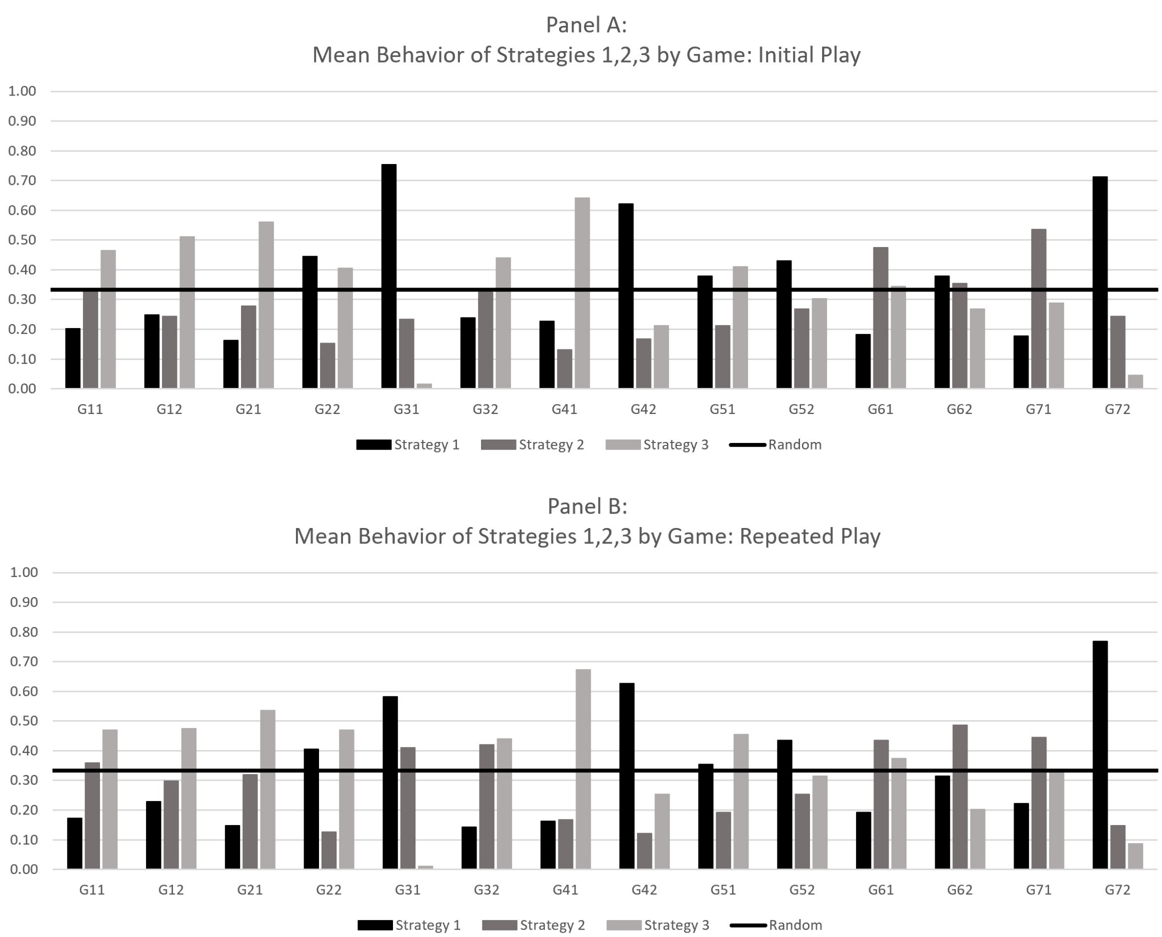


Figure 2: Mean Behavior in Initial and Repeated Play

Figure 2 shows the results for the first (panel A) and second parts (panel B) of the experiment. Clearly, individual behavior is different from random play in both initial and repeated play; otherwise, we would observe that in each game, each of the 3 strategies is played with 1/3 probability (p-values less than 0.001 for both the first and second parts, using a chi-square test against a uniform distribution). Additionally, the mean behavior does not differ significantly between the first and second parts of the experiment, as we cannot reject that the behavior in both scenarios comes from the same distribution (p-value of 0.84 from the two-sample chi-square test that two data samples come from the same distribution), which may suggest that many subjects ignore the provided information on opponent's past behavior and follow the same strategy as in the first part. However, mean behavior can mask important differences with respect to individual heterogeneity. The key task in the next two subsections is to identify the relevant behavioral types that are able to reproduce the behavior in both parts of the experiment.

4.2 Naivete and Sophistication in Initial Play: Type Identification

Using individual data on revealed choices by 198 subjects in 14 different games in the first part of the experiment, we proceed to identify the behavioral type of each subject in initial play. Using a mixture-of-types model with uniform errors, we identify and classify each of the 198 subjects into a behavioral type. The maximum likelihood function is estimated subject by subject. Please see Appendix B for a general description of the maximum likelihood function used to estimate behavioral types and for a particular derivation of the maximum likelihood function for estimating the behavioral types in initial play.

Table 3 shows the estimation results. We allow for different criteria on noise levels or alternatively perfect guesses, from 7 to 11 perfect guesses. We refer to a guess as perfect when a subject's action coincides with a behavioral rule's prediction. Note that, by chance, if individual play was random, any behavioral type that predicts a particular strategy combination across the 14 games would make 4.6 perfect guesses. Therefore, using this value as a benchmark, we consider both less and more stringent identification criteria: no constraints, at least 7 perfect guesses (50% improvement over random), 9 (93% improvement over random) and 11 perfect guesses (139% improvement over random). As expected, a trade-off exists between the number of perfect guesses required for identification and the number of subjects we can properly identify. Nevertheless, remarkably, when imposing 9 perfect guesses (out of 14), which is a high threshold (93% improvement over random), we can identify 93 subjects.[10]

Table 3: Behavioral Type Identification for Initial Play

	Minimum Number of Perfect Guesses			
Model	No Constraints	7	9	11
	(1)	(2)	(3)	(4)
Non-strategic	0.60	0.58	0.53	0.62
A	0.15	0.13	0.10	0.14
IA	0.07	0.06	0.01	0.00
MaxMax	0.11	0.11	0.09	0.05
MaxMin	0.12	0.11	0.11	0.10
L1	0.16	0.17	0.22	0.33
L2	0.36	0.37	0.47	0.38
L3	0.02	0.02	0.00	0.00
NE	0.02	0.02	0.01	0.00
No. of Subjects	198	186	93	21

Notes: The table displays the population frequencies estimated to be consistent with each of the behavioral rules listed in the Model column for different numbers of perfect guesses from all subjects (column 1) to subjects with 7, 9 and 11 (column 4) perfect guesses.

As observed in Table 3, focusing on the overall population, in column 1, 60% of the subjects follow a non-strategic behavioral rule, followed by *L2* (36%), and only a minority of subjects (4%) are identified as sophisticated *L3* and *NE*. Among the non-strategic behavioral types, *L1* and *A* explain most of the behavior, followed by pessimistic and optimistic behavioral rules. These results are roughly consistent with existing results, summarized in Crawford et al. (2013), although we find lower frequencies for *L1* and higher frequencies for *L2*. Furthermore, these conclusions do not change if we move across different columns (criteria over the required perfect guesses). Only when we impose 11 correct guesses, for which we can only identify 21 subjects, do we find considerably more *L1* individuals to the detriment of the optimistic types. However, the overall conclusions remain unchanged: we still find that approximately

[10] Note that we do allow for the existence of Level-*0* type in the estimation. When no model does better than random uniform, the estimated error would be equal to 1, which is interpreted as random uniform play describing best such subject's behavior. We find no subject who is best described by Level-*0*. Also, we do allow for ties between behavioral types, that is, when two behavioral types are equally good in describing a particular subject's action profile over the 14 games. We find 24, 17, 2, and 0 of those cases when we impose no constraints and when we impose 7, 9 and 11 perfect guesses, respectively. When a tie, we assign the subject to the naivest type.

62% of the subject population is identified to follow a non-strategic behavioral rule, followed by *L2* (38%). We cannot reject that the type distribution of the subjects does not depend on the constraints imposed regarding the number of perfect guesses (p-value of 0.12 for the chi-square test), so the estimation results are robust to the criteria on the perfect guesses.[11]

4.3 Naivete and Sophistication in Repeated Play with Information on Past Behavior: Type Identification

As we did for initial play, we identify the behavioral type of subjects by applying a mixture-of-types model with uniform errors to the individual data on revealed choices by the 198 subjects in 14 different games. The general description of the maximum likelihood function can be found in Appendix B, as can a particular derivation of the maximum likelihood function for estimating the behavioral types in repeated play.

Table 4 shows the estimation results. As in the first part of the experiment, we allow for criteria on noise levels or alternatively perfect guesses, from 7 to 11 perfect guesses. Note that, by chance, if individual play was random, any behavioral type that predicts a particular strategy profile would make 4.6 perfect guesses. Therefore, using this value as a benchmark, we allow for less stringent to more stringent identification of behavioral types: no constraints, at least 7 perfect guesses (50% improvement over random), 9 (93% improvement over random) or 11 perfect guesses (139% improvement over random). Again, a trade-off exists between the number of required perfect guesses for identification and the number of subjects we can properly identify. The number of subjects we can cleanly identify is better than that in the first part. When we impose the criterion of 9 perfect guesses, we now identify 144 subjects (73% of the subject population).[12]

[11] Despite not receiving any feedback from game to game, it is still possible that subjects might be learning to be more sophisticated as they play the 14 games. For robustness, we have also estimated the type distribution using only the first half of the 14 games and using only the second half of the 14 games. The estimated type frequency changes slightly, but we do not observe any increase in strategic sophistication from the first to the second half.

[12] As when identifying behavioral models in initial play we do allow for the existence of the Level-*0* type in the estimation. Additionally, we do allow for ties between behavioral types, that is, when two behavioral types are equally good in describing a particular subject's action profile over the 14 games. We find 22, 13, 3, 0 of those cases when we impose no constraints, and when we impose 7, 9 and 11 perfect guesses, respectively. When a tie, we assign the subject to the naivest type.

Table 4: Behavioral Type Identification for Repeated Play

Model	Minimum Number of Perfect Guesses: No Constraints (1)	7 (2)	9 (3)	11 (4)
No-Change	0.36	0.38	0.44	0.50
Adaptive	0.54	0.53	0.51	0.45
$Adaptive_S$	0.17	0.17	0.20	0.28
$Adaptive_A$	0.26	0.26	0.26	0.15
$Adaptive_{IA}$	0.11	0.10	0.05	0.02
Sophisticated	0.07	0.07	0.02	0.04
Sophisticated 2	0.04	0.03	0.03	0.00
No. of Subjects	198	184	125	46

Notes: The table displays the population frequencies estimated to be consistent with each of the behavioral rules listed in the Model column for different numbers of perfect guesses, from all subjects (column 1) to subjects with 7, 9 or 11 (column 4) perfect guesses.

The behavior of 36% of the subjects is best explained by the *No-Change* type, which reflects that an important number of subjects ignore the opponent's past behavior and simply repeat their own past behavior. The most common behavior is adaptive behavior, which was followed by 54% of the subjects, that is, those who best respond to the opponent's past behavior. Among different adaptive learners, the one that is maximizing over the sum of own and opponent's payoffs seems the most frequent one. Finally, very few subjects show sophisticated learning behavior. Consistent with previous findings, these conclusions do not change as we move across different columns. If anything, when the highest threshold of 11 perfect guesses is imposed, the frequency of *No-Change* increases by 14 percentage points to the detriment of both *Adaptive* and *Sophisticated* learning models. As before, we cannot reject that the type distribution of the subjects depends on the constraints imposed regarding the number of perfect guesses (p-value of 0.06 for the chi-square test), so the results are robust to the criteria on perfect guesses.

4.4 Correlation between Naivete and Sophistication in Initial and Repeated Play

We now study the central question of the paper, the correlation between the type identification in initial and repeated play, exploiting the fact that all subjects participated in the same two parts of the experiment. We use a contingency table in which the rows present the behavioral rules in initial play, and the columns present the behavioral rules in repeated play. Therefore, a particular cell in the contingency table shows the proportion of subjects identified as following the behavioral rule in that particular row in initial play who also follow the behavioral rule in that particular column in repeated play. The frequencies across the columns sum to 1 in each row. A positive correlation would show a higher frequency of a naive, non-strategic, behavioral rule in initial play to be using a *No-Change* or less sophisticated rules in repeated play than level-*2* subjects, who would show a higher frequency of learning as adaptive or sophisticated learners. A no-correlation result would show independence in the distributions across different rows. A negative correlation would show that a naive behavioral rule in initial play is using a more sophisticated learning model than a more sophisticated rule in initial play.

As observed in panel A of Table 5, for all 198 subjects, we see little evidence of a positive correlation between naivete and sophistication in initial and repeated play.

On the one hand, 61% of naive subjects in initial play follow an adaptive learning model, and 29% stick to their initial play. *L2* subjects, on the other hand, are almost equally likely to repeat their behavior and to follow an adaptive learning model (49% and 42%, respectively). The few most sophisticated subjects in initial play (*L3* and *Nash*) do show a clear tendency to use adaptive learning models. Focusing on the two most frequent behavioral rules in initial play, naive and *L2*, these numbers show a clear absence of a positive correlation and, if anything, suggestive evidence of a negative correlation in the naivete/sophistication between the initial and repeated play. Panel B shows the equivalent results for a reduced number of subjects when we impose the criterion of 9 perfect guesses. In this case, the subjects show more consistency and therefore a better identification of behavioral rules, although we restrict the sample to 63 subjects. However, the results regarding the correlation in panel B are very similar to those in panel A: 61% of naive subjects in initial play follow an adaptive learning model, and 36% of the *L2* subjects use an adaptive learning model. If anything, the evidence on a negative correlation becomes even stronger. In either panel, we cannot reject that the distributions of the main types are independent across rows (p-values of 0.31 and 0.19, respectively, for the chi-square test).

We therefore conclude that we find no evidence of a positive correlation between naivete and sophistication in initial play and repeated play.

Table 5: Contingency Table

Panel A: No constraints

First Part Model	Second Part Model: *No-change*	*Adaptive*	$Adap_S$	$Adap_A$	$Adap_{IA}$	*Soph*	*Soph 2*	No. of Subjects
Non-strategic	0.29	0.61	0.18	0.29	0.14	0.07	0.03	119
A	0.21	0.72	0.21	0.34	0.17	0.07	0.00	29
IA	0.31	0.46	0.15	0.23	0.08	0.08	0.15	13
$MaxMax$	0.10	0.86	0.24	0.38	0.24	0.00	0.05	21
$MaxMin$	0.54	0.43	0.17	0.13	0.13	0.04	0.00	24
L1	0.31	0.53	0.13	0.31	0.09	0.13	0.03	32
L2	0.49	0.42	0.17	0.21	0.04	0.06	0.04	72
L3	0.25	0.50	0.00	0.50	0.00	0.25	0.00	4
NE	0.33	0.67	0.00	0.00	0.67	0.00	0.00	3
No. of Subjects	72	106	33	51	22	13	7	198

Panel B: Minimum of 9 correct guesses in each part

First Part Model	Second Part Model: *No-Change*	*Adaptive*	$Adap_S$	$Adap_A$	$Adap_{IA}$	*Soph*	*Soph 2*	No. of Subjects
Non-strategic	0.40	0.60	0.19	0.37	0.04	0.00	0.00	27
A	0.17	0.84	0.17	0.67	0.00	0.00	0.00	6
IA	-	-	-	-	-	-	-	0
$MaxMax$	0.00	1.00	0.67	0.33	0.00	0.00	0.00	3
$MaxMin$	0.83	0.17	0.17	0.00	0.00	0.00	0.00	6
L1	0.42	0.58	0.08	0.42	0.08	0.00	0.00	12
L2	0.56	0.36	0.22	0.11	0.03	0.06	0.03	36
L3	-	-	-	-	-	-	-	0
NE	-	-	-	-	-	-	-	0
No. of Subjects	31	29	13	14	2	2	1	63

Notes: The table shows for each of the behavioral rules in initial play (by row) the proportion of subjects identified as following each of the behavioral rules in repeated play. For each row, the proportions across the four columns referring to the four main behavioral rules (*No-Change*, *Adaptive*, *Sophisticated* and *Sophisticated 2*) should sum up to 1. Furthermore, for each row, the proportions across the three adaptive behavioral models ($Adaptive_S$, $Adaptive_A$, $Adaptive_S$) should sum up equal to the value in the column for *Adaptive*.

5 Robustness

One important concern when testing for a correlation between strategic sophistication and naivete in initial and repeated play is that the behavioral type identification is misspecified because some relevant behavioral rules that are relevant to explaining subjects' behavior are not considered. With this concern in mind, we perform three robustness tests. First, we repeat the estimation with elicited behavior in the first part including several alternative behavioral rules in addition to those we already considered. Second, we perform an omitted type specification test to alternatively confirm whether we obtain our result due to the omission of one or many relevant behavioral rules. Finally, as we find a high number of *No-Change* type, almost 60% of the subjects, we perform an additional analysis replacing the *No-Change* type with all the behavioral rules we considered in the first part.

5.1 Addition of Alternative Behavioral Rules in Initial Play

We consider 4 alternative behavioral types for the initial play in addition to the 8 we described in Section 3.1. All four types could be considered to be variations of *L1*, where we alter the belief about opponent's behavior. Given that we consider it to be plausible that subjects follow some simple non-strategic rules, it is also plausible that some subjects think in the same way. Consequently, we consider *L1* as best responding to each of the other non-strategic rules we initially included, that is, $L1_{A}$, $L1_{IA}$, $L1_{MaxMax}$ and $L1_{MaxMin}$. Note that these alternative behavioral rules are clearly strategic and closer in spirit to *L2* in terms of strategic sophistication, as they predict a particular opponent's strategy and best response to that strategy. Additionally, as shown in Table A2 in the Appendix, these additional behavioral types show good separation from the types we initially considered.

As shown in Table A3 in the Appendix, the alternative models appear to show some relevance, although they do not alter the identified type distribution substantially. First, as expected, the new alternative behavioral rules steal frequency mostly from *L2* and the non-strategic types (mostly *A*). The additional behavioral model that appears to be the most relevant is $L1_{MaxMax}$, which is followed by 9% of subjects. The contingency table displayed in Table 6 shows that subjects following these alternative models are best explained by *No-Change* and *Adaptive*, and only a minority are best explained by *Sophisticated* in repeated play. In summary, the consideration of additional alternative behavioral rules to explain initial play does not alter the main results: we find no evidence of a positive correlation between naivete and sophistication in initial and repeated play.

Table 6: Contingency Table with Additional Alternative Behavioral Rules: All Subjects

First Part Model	Second Part Model: *No-Change*	*Adaptive*	$Adap_S$	$Adap_A$	$Adap_{IA}$	*Soph*	*Soph 2*	No. of Subjects
Non-strategic	0.31	0.60	0.17	0.28	0.15	0.06	0.03	104
A	0.25	0.70	0.20	0.30	0.20	0.05	0.00	20
IA	0.33	0.50	0.17	0.25	0.08	0.08	0.08	12
$MaxMax$	0.11	0.84	0.21	0.37	0.26	0.00	0.05	19
$MaxMin$	0.52	0.43	0.17	0.13	0.13	0.04	0.00	23
L1	0.30	0.56	0.13	0.33	0.10	0.10	0.03	30
Alternative Models	0.37	0.43	0.17	0.24	0.02	0.12	0.07	41
$L1_A$	0.27	0.63	0.18	0.36	0.09	0.00	0.09	11
$L1_{IA}$	0.43	0.43	0.14	0.29	0.00	0.14	0.00	7
$L1_{MaxMax}$	0.41	0.36	0.18	0.18	0.00	0.14	0.09	22
$L1_{MaxMin}$	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1
L2	0.48	0.46	0.16	0.24	0.06	0.04	0.02	50
L3	-	-	-	-	-	-	-	0
NE	0.33	0.67	0.00	0.67	0.00	0.00	0.00	3
No. of Subjects	72	106	33	51	22	13	7	198

Notes: The table shows for each of the behavioral rules in initial play (by row) the proportion of subjects identified as following each of the behavioral rules in repeated play. For each row, the proportions across the four columns referring to the four main behavioral rules (*No-Change*, *Adaptive*, *Sophisticated* and *Sophisticated 2*) should sum up to 1. Furthermore, for each row, the proportions across the three adaptive behavioral models ($Adaptive_S$, $Adaptive_A$, $Adaptive_S$) should sum up equal to the value in the column for *Adaptive*.

5.2 Specification Test: Omitted Types

Similar in spirit to the previous robustness test, we also perform an omitted type specification test (as in Costa-Gomes and Crawford, 2006) to rule out the possibility that we did not consider relevant models.

In this test, instead of proposing alternative behavioral models, we let the actual subject behavior in our sample inform us of potential alternative rules. If we left out a rule that actually complies with subjects' behavior, we would expect that some of the subjects behave similarly to this rule. Therefore, we consider the observed behavior as potential new rules in the following manner. In addition to all 12 behavioral rules considered in the previous section, we add each subject's actual behavior as an additional behavioral rule, one subject at a time, and re-estimate the mixture-of-types model as many times as the number of subjects in our population, that is, 198 times. While conducting this exercise, we check whether the added subject's behavioral rule is able to explain other subjects' behavior *better* than the existing 12 models and whether the rule can attract sufficient relevance, where we impose a threshold of 15% of the population frequency.

We find three such subjects (subject numbers 31, 85, and 86). What strategies are these subjects following? First, we check for similarity of these subjects' behavior (or alternatively, separation). These subjects appear to reflect the same type of behavior, as they show very little separation (0.21 between the behavior of subject 31 and subject 85, 0.14 between the behavior of subject 31 and subject 86, and 0.36 between the behavior of 85 and 86). Second, we check their separation from other existing behavioral rules, as shown in Table A4 in the Appendix. All three behavioral rules are well separated from all other considered rules, with the exception of *L2*, which shows a separation equal to or less than 0.43. Third, consistent with this finding, we also observe that when we consider these alternative models in the mixture-of-types model estimation, the behavioral rule that loses the most frequency is indeed *L2*, as shown by estimations in Table A5. Finally, we directly consider the actions of these subjects and find that their behavior is mostly consistent with *L2*, but in a few games, it mimics *L1*.[13]. In particular, the strategy profiles of subjects 85 and 31 diverge from *L2* or *L1* behavior in only two decisions and that of subject 86 diverges in only three decisions.

We conclude that these subjects show some variation from the existing *L2* behavioral type; however, none of them obtains a population frequency higher than that of *L2* when incorporated into the estimation together, as shown in Table A5, or one by one.

Does the result of the correlation between sophistication and naivete between initial and repeated play change when these new empirically motivated behavioral rules are considered? Table 7 shows that subjects following these alternative models are best explained by *No-Change* and by adaptive learners, with similar proportions as those in Table 5. Therefore, we again conclude that we find no evidence for a positive correlation

[13]In particular, the strategy profile of subject 31 is 3 1 3 3 3 1 3 3 1 2 1 2 3 1; the strategy profile of subject 85 is 3 1 3 3 3 1 1 3 1 2 3 3 3 1; and the strategy profile of subject 86 is 3 1 3 1 3 1 3 2 1 2 1 2 3 1.

between naivete and sophistication in initial and repeated play.

Table 7: Contingency Table with the Addition of Three Subjects' Behavioral Rules: All Subjects

	Second Part Model							
First Part Model	*No-Change*	*Adaptive*				*Soph*	*Soph 2*	No. of Subjects
			$Adap_S$	$Adap_A$	$Adap_{IA}$			
Non-strategic	0.32	0.59	0.16	0.28	0.15	0.05	0.03	74
A	0.36	0.54	0.09	0.45	0.00	0.09	0.00	11
IA	0.33	0.51	0.17	0.17	0.17	0.00	0.17	6
$MaxMax$	0.14	0.79	0.21	0.29	0.29	0.00	0.07	14
$MaxMin$	0.50	0.45	0.15	0.15	0.15	0.05	0.00	20
L1	0.26	0.65	0.17	0.35	0.13	0.09	0.00	23
Alternative Models	0.29	0.45	0.19	0.23	0.03	0.16	0.10	31
$L1_A$	0.00	0.83	0.33	0.33	0.17	0.00	0.17	6
$L1_{IA}$	0.33	0.50	0.17	0.33	0.00	0.17	0.00	6
$L1_{MaxMax}$	0.39	0.34	0.17	0.17	0.00	0.17	0.11	18
$L1_{MaxMin}$	0.00	0.00	0.00	0.00	0.00	1.00	0.00	1
Subject 31	0.56	0.32	0.00	0.13	0.19	0.06	0.06	16
Subject 85	0.30	0.65	0.20	0.30	0.15	0.05	0.00	20
Subject 86	0.38	0.58	0.19	0.29	0.10	0.05	0.00	21
L2	0.44	0.50	0.21	0.26	0.03	0.03	0.03	34
L3	-	-	-	-	-	-	-	0
NE	0.50	0.50	0.00	0.00	0.50	0.00	0.00	2
No. of Subjects	72	106	33	51	22	13	7	198

Notes: The table shows for each of the behavioral rules in initial play (by row) the proportion of subjects identified as following each of the behavioral rules in repeated play. For each row, the proportions across the four columns referring to the four main behavioral rules (*No-Change*, *Adaptive*, *Sophisticated* and *Sophisticated 2*) should sum up to 1. Furthermore, for each row, the proportions across the three adaptive behavioral models ($Adaptive_S$, $Adaptive_A$, $Adaptive_S$) should sum up equal to the value in the column for *Adaptive*.

5.3 Specification Test: Replacement of *No-Change* type

We found that the large majority of subjects, 60% of them, followed the simplest *No-Change* behavioral type in repeated play, such that their behavior is best explained by simply taking exactly the same strategy as the one they did in the first part. We therefore question whether our results would significantly change if we replaced the *No-Change* type by all the behavioral rules we considered in the first stage and added the more sophisticated learning models such as *Adaptive*, *Sophisticated* and *Sophisticated 2*.

Table A6 in the Appendix shows these results. The behavioral rules from the first stage maintain relative frequencies similar to the original estimation see Table 3. The frequency of the models from the second stage decrease slightly as they have to compete with more alternative explanations, but more importantly they remain relevant and show the same frequency ordering. *Adaptive* is the most frequent non-naive model, followed by *Sophisticated* and *Sophisticated 2*, which show a much lower frequency, as before.

Table 8: Contingency Table with *No-Change* type replaced by Behavioral Models in Initial Responses: All Subjects

	Second Part Model							
First Part Model	Models of initial play	*Adaptive*				*Soph*	*Soph 2*	No. of Subjects
			$Adap_S$	$Adap_A$	$Adap_{IA}$			
Non-strategic	0.46	0.51	0.13	0.28	0.10	0.03	0.01	90
A	0.24	0.76	0.21	0.38	0.17	0.00	0.00	29
IA	0.38	0.61	0.15	0.38	0.08	0.00	0.00	13
$MaxMax$	0.38	0.57	0.10	0.33	0.14	0.00	0.05	21
$MaxMin$	0.50	0.46	0.17	0.21	0.08	0.04	0.00	24
L1	0.72	0.22	0.03	0.16	0.03	0.06	0.00	32
L2	0.44	0.44	0.21	0.19	0.04	0.06	0.06	72
L3	0.25	0.75	0.00	0.75	0.00	0.00	0.00	4
NE	0.67	0.33	0.00	0.00	0.33	0.00	0.00	3
No. of Subjects	90	96	30	50	16	7	8	198

Notes: The table shows for each of the behavioral rules in initial play (by row) the proportion of subjects identified as following each of the behavioral rules in repeated play. For each row, the proportions across the four columns referring to the four main behavioral rules (*No-Change*, *Adaptive*, *Sophisticated* and *Sophisticated 2*) should sum up to 1. Furthermore, for each row, the proportions across the three adaptive behavioral models ($Adaptive_S$, $Adaptive_A$, $Adaptive_S$) should sum up equal to the value in the column for *Adaptive*.

We can finally reproduce the contingency table replacing the *No-Change* type by all models that we considered in initial play, as shown in Table 8. The two most important models, *L2* and the non-strategic ones, show a very similar distribution over the models considered in repeated play. In other words, the non-strategic models and the *L2* behavioral types show very similar frequencies of using initial play models as before and an adaptive learning model, showing one more time that the naivete and sophistication in the first play show little correlation with naivete and sophistication in repeated play.

The full table, where we disaggregate the *No-Change* into the different behavioral types included in part 1, is shown in Table A7 in the Appendix. Sometimes, subjects whose behavior is best described by *No-Change* change type between the first and the second parts. However, if a subject is identified as *No-Change*, it is more likely that this subject is identified as using exactly the same behavioral type as in the first part. For example, subjects identified as following *L2* behavioral type are identified mostly into two behavioral rules in the second part: either *L2* type or *Adaptive* (either $Adaptive_S$ or $Adaptive_A$).

More importantly, replacing *No-Change* does not change the estimated frequency of *Adaptive* nor the correlation of sophistication between initial and repeated play. We therefore conclude that the results are robust to replacing the *No-Change* type with models considered in initial play.

6 Discussion

In this paper, we have explored the relationship between the strategic sophistication and naivete of models in initial and repeated play. Is a strategically naive player in initial play more likely than a more sophisticated player to use a naive model in repeated play? We use an experimental design and mixture-of-types model econometric estimation to answer this empirically motivated research question.

Consistent with previous findings, we find that the Nash equilibrium is not well suited to explain the initial responses of individuals. The non-equilibrium rules that best explain individual behavior appear to be level-*2*, level-*1* and altruistic type thinking. Additionally, consistent with previous findings, adaptive behavior appears to be the most common learning model, although an important number of individuals simply repeat the previously used strategy. Addressing the central question, exploiting the within-subject design, we find that the behavior in repeated play is not positively correlated with the behavior in initial play.

The main result of our paper is reminiscent of the results of Costa-Gomes and Weizsacker (2008) and Knoepfle et al. (2009). The former found an inconsistency between the behavior shown by actions and elicited beliefs regarding opponents' expected behavior. The latter found that eye-tracking results favor much more sophisticated learning than do actual decision data, again finding an inconsistency between the two. It could indeed be the case that, similar to actions and beliefs or actions and eye-tracking, individuals treat initial and repeated play as different and/or independent tasks.

References

[1] Alaoui, L., and Penta, A. (2016). "Endogenous depth of reasoning". *The Review of Economic Studies*, 83(4), 1297-1333.

[2] Brañas-Garza, P., Espinosa, M. P., and P. Rey-Biel (2011). "Travelers types". *Journal of Economic Behavior and Organization*, 78(1-2), 25-36.

[3] Camerer, C., and T. H. Ho (1999). "Experience-weighted attraction learning in normal form games." *Econometrica*, 67(4), 827-874.

[4] Camerer, Colin F., T. H. Ho, and J. K. Chong (2002). "Sophisticated experience-weighted attraction learning and strategic teaching in repeated games." *Journal of Economic theory* 104.1: 137-188.

[5] Camerer, C. F., Ho, T. H., and J. K. Chong (2004). "A cognitive hierarchy model of games". *The Quarterly Journal of Economics*, 119(3), 861-898.

[6] Cheung, Y. W., and Friedman, D. (1997). "Individual learning in normal form games: Some laboratory results". *Games and economic behavior*, 19(1), 46-76.

[7] Conlisk, J. (1993a). "Adaptation in games: Two solutions to the Crawford puzzle". *Journal of Economic Behavior and Organization*, 22(1), 25-50.

[8] Conlisk, J. (1993b). "Adaptive tactics in games: Further solutions to the Crawford puzzle". *Journal of Economic Behavior and Organization*, 22(1), 51-68.

[9] Costa-Gomes, M. A., V. P. Crawford, and B. Broseta (2001). "Cognition and behavior in normal-form games: An experimental study," *Econometrica*, 69: 1193-1235.

[10] Costa-Gomes, M. A., and V. P. Crawford (2006). "Cognition and behavior in two-person guessing games: An experimental study." *The American Economic Review*, 96.5: 1737-1768.

[11] Costa-Gomes, M. A., and G. Weizsacker (2008). "Stated beliefs and play in normal-form games". *The Review of Economic Studies*, 75(3), 729-762.

[12] Crawford, V.P., M. A. Costa-Gomes and N. Iriberri (2013). “Structural Models of Non-equilibrium Strategic Thinking:Theory, Evidence, and Applications,” *Journal of Economic Literature*, 51:1: 5-62.

[13] Eckel, Catherine C. and P. J. Grossman (2002). “Sex differences and statistical stereotyping in attitudes toward financial risk.” *Evolution and Human Behavior*, 23(4), 281295.

[14] Erev, I., and A. E. Roth (1998). “Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria”. *The American Economic Review*, 848-881.

[15] Eyster, E. (2003). “Rationalising the Past: A Taste for Consistency”. Mimeo.

[16] Falk, A., and Zimmermann, F. (2011). “Preferences for consistency”. Mimeo.

[17] Feltovich, N. (2000). “Reinforcementbased vs. Beliefbased Learning Models in Experimental Asymmetricinformation Games”. *Econometrica*, 68(3), 605-641.

[18] Fischbacher, U. (2007). “z-Tree: Zurich toolbox for ready-made economic experiments.” *Experimental Economics*, 10(2), 171-178.

[19] Frederick, S. (2005). “Cognitive reflection and decision making.” *Journal of Economic perspectives*, 19(4), 25-42.

[20] Fudenberg, D., and D. Levine. (1998a). “Learning in games.” *European Economic Review* 42.3-5: 631-639.

[21] Fudenberg, D., and D. Levine. (1998b). Theory of learning in games. Cambridge, MA: MIT Press.

[22] García-Pola, B., Iriberri, N. and J. Kovářík (2020). “Non-Equilibrium Play in Centipede Games.” *Games and Economic Behavior* 120: 391433.

[23] Gill, D., and Prowse, V. (2016). “Cognitive ability, character skills, and learning to play equilibrium: A level-k analysis.” *Journal of Political Economy*, 124(6), 1619-1676.

[24] Goeree, J. K., and C. A. Holt (2001). "Ten little treasures of game theory and ten intuitive contradictions." *The American Economic Review* 91.5: 1402-1422.

[25] Greiner, B. (2015). "Subject pool recruitment procedures: organizing experiments with ORSEE". *Journal of the Economic Science Association*, 1(1), 114-125.

[26] Ho, T.H., Park, S.E., and X. Su (Forthcoming). "A Bayesian Level-k Model in n-Person Games." *Management Science*.

[27] Iriberri, N., and P. Rey-Biel (2013). "Elicited beliefs and social information in modified dictator games: what do dictators believe other dictators do". *Quantitative Economics* 4: 515-547.

[28] Knoepfle, D. T., Wang, J. T. Y., and C. F. Camerer (2009). "Studying learning in games using eye-tracking". *Journal of the European Economic Association*, 7(2-3), 388-398.

[29] Kovářík, J., Mengel, F., and Romero, J. G. (2018). "Learning in network games". *Quantitative Economics*, 9(1), 85-139.

[30] McKelvey, R. D., and T. R. Palfrey (1995). "Quantal Response Equilibrium for Normal-Form Games". *Games and Economic Behavior* 10: 6–38.

[31] Milgrom, P., and J. Roberts (1991). "Adaptive and sophisticated learning in normal form games". *Games and economic Behavior*, 3(1), 82-100.

[32] Nagel, R. (1995). "Unraveling in guessing games: An experimental study." *The American Economic Review* 85.5: 1313-1326.

[33] Rey-Biel, P. (2009). "Equilibrium play and best response to (stated) beliefs in normal form games". *Games and Economic Behavior*, 65(2), 572-585.

[34] Selten, R. (1991). "Anticipatory learning in two-person games." In Game equilibrium models I (pp. 98-154). Springer, Berlin, Heidelberg.

[35] Sobel, J. (2000). "Economists' models of learning." *Journal of Economic Theory* 94.2: 241-261.

[36] Stahl, D. O., and Wilson, P.W. (1994). “Experimental evidence on players’ models of other players.” *Journal of Economic Behavior and Organization* 25.3: 309-327.

[37] Stahl, D. O., and Wilson, P.W. (1995). “On players’ models of other players: Theory and experimental evidence.” *Games and Economic Behavior* 10.1: 218-254.

[38] Stahl, D. O., (2003). “Sophisticated Learning and Learning Sophistication.” Mimeo.

[39] Toplak, M. E., West, R. F., and K. E. Stanovich (2014). “Assessing miserly information processing: An expansion of the Cognitive Reflection Test.” *Thinking and Reasoning*, 20(2), 147-168.

A Additional Tables And Figures

Table A1: Summary of Socio-Demographic Variables of the Subject Population

Variables	Mean Values	Stand. Dev.
Men	0.41	
Age	21.73	2.99
Spanish	0.87	
University Entry Grade (out of 10)	6.85	1.16
Distribution over Field of Study:		
Social Science	0.77	
Applied Science	0.17	
Natural Science	0.04	
Distribution over risk choices:		
1.5€with 0.50 or 1.5€with 0.50	0.31	
1.3€with 0.50 or 1.8€with 0.50	0.11	
1.1€with 0.50 or 2.1€with 0.50	0.26	
0.9€with 0.50 or 2.4€with 0.50	0.07	
0.7€with 0.50 or 2.7€with 0.50	0.04	
0.6€with 0.50 or 2.8€with 0.50	0.04	
0.4€with 0.50 or 2.9€with 0.50	0.02	
0€with 0.50 or 3€with 0.50	0.16	
Cognitive reflection test:		
Percent correct in cognitive reflection test: Q1	0.28	
Percent correct in cognitive reflection test: Q2	0.17	
Percent correct in cognitive reflection test: Q3	0.41	

Notes: *Men* takes a value of 1 if the subject is male. *Age* reflects the age in years. *Spanish* takes a value of 1 if the subject is Spanish. *University Entry Grade* is normalized to a grade out of 10. *Social Science*, *Applied Science* and *Natural Science* take values of 1 if the subject is studying a social, applied or natural science. Risk Choice was elicited via Eckel and Grossman (2002), where choices are ordered from safest to riskiest. Finally, the cognitive reflection test includes questions from Toplak et al. (2014) designed to avoid the possibility that the original test from Frederick (2005) is already known by the subjects. The questions are as follows: 1. If John can drink one barrel of water in 6 days, and Mary can drink one barrel of water in 12 days, how long would it take them to drink one barrel of water together? (correct answer 4 days; intuitive answer 9); 2. Jerry received both the 15th highest and the 15th lowest mark in the class. How many students are in the class? (correct answer 29 students; intuitive answer 30); 3. A man buys a pig for \$60, sells it for \$70, buys it back for \$80, and sells it finally for \$90. How much has he made? (correct answer \$20; intuitive answer \$10).

Table A2: Separation of Different Behavioral Rules with Additional Alternative Behavioral Models

	A	*IA*	*MaxMax*	*MaxMin*	*L1*	$L1_A$	$L1_{IA}$	$L1_{MaxMax}$	$L1_{MaxMin}$	*L2*	*L3*	*NE*
A	0.00											
IA	0.60	0.00										
MaxMax	0.46	0.62	0.00									
MaxMin	0.71	0.65	0.57	0.00								
L1	0.57	0.76	0.50	0.79	0.00							
$L1_A$	0.25	0.62	0.36	0.64	0.64	0.00						
$L1_{IA}$	0.64	0.98	0.64	0.71	0.50	0.57	0.00					
$L1_{MaxMax}$	0.75	0.73	0.64	0.57	0.79	0.50	0.64	0.00				
$L1_{MaxMin}$	0.71	0.65	0.93	0.64	0.71	0.71	0.64	0.57	0.00			
L2	0.75	0.58	0.57	0.64	0.71	0.57	0.79	0.5	0.79	0.00		
L3	0.86	0.80	0.86	0.64	0.79	0.71	0.43	0.50	0.50	0.71	0.00	
NE	0.57	0.51	0.86	0.64	0.79	0.57	0.50	0.86	0.57	0.79	0.57	0.00

Notes: The table reports the proportion of strategies across all 14 games in which the different behavioral models predict different strategies. The minimum possible separation value is 0, which occurs when the two models prescribe the same strategy in all 14 games, and the maximum possible separation value is 1, which occurs when the two models predict a different strategy in each of the 14 games.

Table A3: Behavioral Type Identification for Initial Play: Additional Behavioral Types

	Minimum Number of Perfect Guesses							
	No constraints		7		9		11	
	Main	Alt.	Main	Alt.	Main	Alt.	Main	Alt.
A	0.15	0.10	0.13	0.10	0.10	0.07	0.14	0.11
IA	0.07	0.06	0.06	0.05	0.01	0.01	0.00	0.00
$MaxMax$	0.11	0.10	0.11	0.10	0.09	0.07	0.05	0.04
$MaxMin$	0.12	0.12	0.11	0.11	0.11	0.09	0.10	0.07
$L1$	0.16	0.15	0.17	0.16	0.22	0.18	0.33	0.25
$L1_A$		0.06		0.06		0.04		0.00
$L1_{IA}$		0.04		0.04		0.03		0.04
$L1_{MaxMax}$		0.11		0.11		0.15		0.21
$L1_{MaxMin}$		0.01		0.01		0.01		0.00
$L2$	0.36	0.25	0.37	0.26	0.47	0.35	0.38	0.29
$L3$	0.02	0.00	0.02	0.00	0.00	0.00	0.00	0.00
NE	0.02	0.02	0.02	0.02	0.01	0.01	0.00	0.00
No. of Subjects	198	198	186	189	93	113	21	28

Notes: The table displays the population frequencies estimated for the main specification shown in Table 3 and when adding alternative models in initial play.

Table A4: Separation of the Three Relevant Subjects' Behavior from other Behavioral Models

	A	IA	$MaxMax$	$MaxMin$	$L1$	$L1_A$	$L1_{IA}$	$L1_{MaxMax}$	$L1_{MaxMin}$	$L2$	$L3$	NE
Subject 31	0.57	0.58	0.71	0.71	0.50	0.57	0.57	0.57	0.64	0.36	0.71	0.57
Subject 85	0.54	0.65	0.50	0.71	0.57	0.50	0.50	0.64	0.79	0.36	0.79	0.64
Subject 86	0.57	0.55	0.64	0.71	0.50	0.57	0.57	0.50	0.64	0.43	0.79	0.71

Notes: The table reports the proportion of strategies across all 14 games in which the three subjects' behavioral models predict different strategies from the rest of the considered models. The minimum possible separation value is 0, which occurs when the two models prescribe the same strategy in all 14 games, and the maximum possible separation value is 1, which occurs when two models predict a different strategy in each of the 14 games.

Table A5: Behavioral Type Identification for Initial Play: Additional Behavioral Types

	Minimum Number of Perfect Guesses							
	No constraints		7		9		11	
	Main	Alt.	Main	Alt.	Main	Alt.	Main	Alt.
A	0.10	0.06	0.10	0.06	0.07	0.06	0.11	0.07
IA	0.06	0.03	0.05	0.03	0.01	0.01	0.00	0.00
$MaxMax$	0.10	0.07	0.10	0.07	0.07	0.05	0.04	0.02
$MaxMin$	0.12	0.10	0.11	0.09	0.09	0.07	0.07	0.04
$L1$	0.15	0.12	0.16	0.12	0.18	0.13	0.25	0.16
$L1_A$	0.06	0.03	0.06	0.03	0.04	0.01	0.00	0.00
$L1_{IA}$	0.04	0.03	0.04	0.03	0.03	0.02	0.04	0.02
$L1_{MaxMax}$	0.11	0.09	0.11	0.09	0.15	0.11	0.21	0.13
$L1_{MaxMin}$	0.01	0.01	0.01	0.01	0.01	0.01	0.00	0.00
$L2$	0.25	0.17	0.26	0.017	0.35	0.21	0.29	0.18
$L3$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
NE	0.02	0.01	0.02	0.01	0.01	0.01	0.00	0.00
Subject 31		0.08		0.08		0.09		0.07
Subject 85		0.10		0.10		0.11		0.16
Subject 86		0.11		0.11		0.12		0.16
No. of Subjects	198	198	189	195	113	142	28	45

Notes: The table displays the population frequencies estimated for the main specification shown in Table 3 and when adding alternative models in initial play.

Table A6: Behavioral Type Identification for Repeated Play: *No-Change* replaced by Behavioral Rules from Initial Play

	Minimum Number of Perfect Guesses							
	No constraints		7		9		11	
	Main	Alt.	Main	Alt.	Main	Alt.	Main	Alt.
No-Change	0.36		0.38		0.44		0.50	
A		0.04		0.03		0.04		0.03
IA		0.02		0.02		0.01		0.00
$MaxMax$		0.05		0.05		0.05		0.03
$MaxMin$		0.07		0.07		0.06		0.06
L1		0.10		0.09		0.09		0.15
L2		0.15		0.15		0.15		0.03
L3		0.01		0.02		0.01		0.00
NE		0.01		0.01		0.02		0.00
Adaptive	0.54	0.48	0.53	0.50	0.51	0.53	0.45	0.65
$Adaptive_S$	0.17	0.15	0.17	0.16	0.20	0.22	0.28	0.28
$Adaptive_A$	0.26	0.25	0.26	0.26	0.26	0.28	0.15	0.24
$Adaptive_{IA}$	0.11	0.08	0.10	0.08	0.05	0.03	0.02	0.03
Sophisticated	0.07	0.04	0.07	0.04	0.02	0.02	0.04	0.06
Sophisticated 2	0.04	0.03	0.03	0.03	0.03	0.04	0.00	0.00
No. of Subjects	198	198	184	191	125	129	45	34

Notes: The table displays the population frequencies estimated for the main specification shown in Table 4 and when replacing the *No-Change* type by all the models included in initial play as in Table 3.

Table A7: Full Contingency Table with No-Change type replaced by Behavioral Models in Initial Responses: All Subjects

First Part Model	Second Part Model: *A*	*IA*	*MaxMax*	*MaxMin*	*L1*	*L2*	*L3*	*NE*	*Adapt*				*Sophis*	*Sophis 2*	No. of Subjects
										$Adap_S$	$Adap_A$	$Adap_{IA}$			
Non-strategic	0.04	0.03	0.04	0.08	0.10	0.13	0.03	0.01	0.51	0.13	0.28	0.10	0.03	0.01	119
A	0.07	0.00	0.07	0.00	0.00	0.10	0.00	0.00	0.76	0.21	0.38	0.17	0.00	0.00	29
IA	0.00	0.08	0.00	0.08	0.08	0.00	0.15	0.00	0.31	0.15	0.38	0.08	0.00	0.00	13
MaxMax	0.00	0.05	0.05	0.14	0.10	0.05	0.00	0.00	0.57	0.10	0.33	0.14	0.00	0.05	21
MaxMin	0.00	0.04	0.00	0.21	0.04	0.21	0.00	0.00	0.46	0.17	0.21	0.08	0.04	0.00	24
L1	0.09	0.00	0.06	0.03	0.25	0.19	0.06	0.03	0.22	0.03	0.16	0.03	0.06	0.00	32
L2	0.04	0.01	0.06	0.06	0.08	0.18	0.00	0.01	0.44	0.21	0.19	0.04	0.06	0.06	72
L3	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.75	0.00	0.75	0.00	0.00	0.00	4
NE	0.00	0.00	0.00	0.00	0.33	0.33	0.00	0.00	0.33	0.00	0.00	0.33	0.00	0.00	3
No. of Subjects	8	4	10	14	19	29	4	2	96	30	50	16	7	5	198

Notes: The table shows for each of the behavioral rules in initial play (by row) the proportion of subjects identified as following each of the behavioral rules in repeated play. For each row, the proportions across columns referring to the 11 main behavioral rules (*A*, *IA*, *MaxMax*, *MaxMin*, *L1*, *L2*, *L3*, *NE*, *Adaptive*, *Sophisticated* and *Sophisticated 2*) should sum up to 1. Furthermore, for each row, the proportions across the three adaptive behavioral models ($Adaptive_S$, $Adaptive_A$, $Adaptive_S$) should sum up equal to the value in the column for *Adaptive*.

B Mixture-of-types Likelihood Function

We assume that a subject i employing rule k makes type-k's decision with probability $(1-\varepsilon_i)$ but makes a mistake with probability $\varepsilon_i \in [0,1]$. In such a case, she plays each of the three available strategies uniformly at random. As in most mixture-of-types model applications, we assume that the errors are identically and independently distributed across games and that the errors are subject-specific (as in, for example, Iriberri and Rey-Biel, 2013). The first assumption facilitates the statistical treatment of the data, while the second considers that some subjects may be noisier and thus make more errors than others.

The likelihood of a particular individual of a particular type can be constructed as follows. Let $P_k^{g,j}$ be type-k's predicted choice probability for strategy j in game g. Some rules may predict more than one strategy in a particular game. This characteristic is reflected in the vector $P_k^g = (P_k^{g,1}, P_k^{g,2}, P_k^{g,3})$ with $\sum_j P_k^{g,j} = 1$.

For each individual in each game, we observe the choice and whether it is consistent with k. Let $x_i^{g,j} = 1$ if strategy j is chosen by subject i in game g in the experiment and $x_i^{g,j} = 0$ otherwise. The likelihood of observing a sample $x_i = (x_i^{g,j})_{g,j}$ given type k and subject i is then

$$L_i^k(\varepsilon_i|x_i) = \prod\nolimits_g \prod\nolimits_j \left[(1-\varepsilon_i)P_k^{g,j} + \frac{\varepsilon_i}{3}\right]^{x_i^{g,j}} . \tag{1}$$

Finally, the likelihood function is given by the sum of all behavioral types that are considered.

$$L_i(\varepsilon_i|x_i) = \sum\nolimits_k p_i L_i^k(\varepsilon_i|x_i) \tag{2}$$

p_i takes a value of 1 for the behavioral type k that best explains the individual behavior and 0 for the rest of the considered behavioral types.

For explaining initial play, we consider $K = 8$ behavioral types or models: *A*, *IA*, *MaxMax*, *MaxMin*, *L1*, *L2*, *L3* and *NE*, and use their revealed actions as input data. To explain repeated play with provided information on past actions, we consider $K = 6$ different behavioral types: *No-Change*, *Adaptive*$_S$, *Adaptive*$_A$, *Adaptive*$_I$*A*, *Sophisticated* and *Sophisticated 2*, and use their revealed actions and observed own and opponent's past action as input data.

C Translation of Instructions

The original instructions were in Spanish. Here we provide a translation of the instructions into English.

THANK YOU FOR PARTICIPATING IN OUR EXPERIMENT!

We will now start the experiment. From now on, you are not allowed to speak, look at what other participants do or walk around the room. Please turn off your phone. If you have any questions or need help, raise your hand and one of the researchers will talk with you. Please, do not write on these instructions. If you do not follow these rules, YOU WILL BE ASKED TO LEAVE THE EXPERIMENT, AND NO PAYMENT WILL BE GIVEN TO YOU. Thank you.

The university and the research projects have provided the funds for the realization of this experiment. You will receive 3 Euros for having arrived on time. Additionally, if you follow the instructions correctly, you have the possibility to earn more money. This is a group experiment. The amount you can earn depends on your decisions, the decisions of other participants, and chance. Different participants can earn different amounts.

No participant will be able to identify another by their decisions or by their profits in the experiment. The researchers will be able to observe the profits of each participant at the end of the experiment, but we will not associate the decisions you have made with the identity of any participant.

EARNINGS:

During the experiment you can earn experimental points. At the end, each experimental point will be exchanged for Euros, and 1 experimental point is worth exactly 0.5 Euros. Everything you win will be paid in cash in a strictly private way at the end of the experimental session.

Your final earnings will be the sum of the 3 Euros you receive for participating plus what you earn during the experiment.

Each experimental point equals 50 cents, so 2 experimental points equals 1 Euro ($2x0.5 = 1$ Euro).

If, for example, you earn a total of 20 experimental points, you will receive a total of 13 Euros (3 Euros as payment for participation and 10 Euros from the conversion of the 20 experimental points to Euros).

If, for example, you earn 4 experimental points, you will obtain 5 Euros ($4x0.5 = 2$ and $2 + 3 = 5$).

If, for example, you earn 44 experimental points, you will obtain 25 Euros ($44x0.5 = 22$ and $22 + 3 = 25$).

PARTS OF THE EXPERIMENT:

The experiment consists of two parts. You will participate by operating a computer. In the first part, there will be 14 rounds where you will make 14 decisions. In the second part, there will also be 14 rounds where you will make 14 decisions. At the end of the experiment, when you have completed the two parts of the experiment, the computer will randomly choose two of the 28 rounds, and you will be paid for the experimental points you received in those two rounds chosen at random, plus the 3 Euros for participating.

Before beginning each part of the experiment, we will explain in detail what kinds of decisions you can make and how you can obtain experimental points.

When we are all ready, we will start the first part of the experiment by explaining the instructions of the first part of the experiment in detail.

FIRST PART OF THE EXPERIMENT:

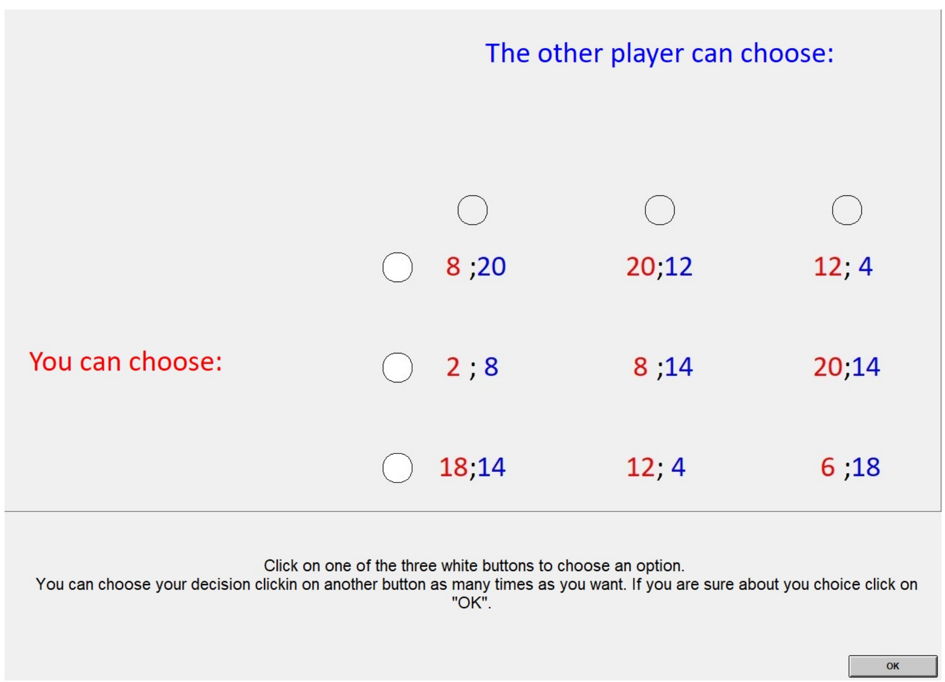
The first part of the experiment consists of 14 rounds. In each of the 14 rounds, you will be paired with a participant chosen at random from this session. The other participant will be different in each of the rounds, so you will never be paired with the same participant more than once. From now on, we will refer to you as "You" and the other participant as "other participant".

In each round, you will have to make a decision by choosing among three possible options. Each decision will be presented in the form of a table similar to the one below (but with different values). You will see the corresponding table each time you have to choose an option. Each row of the table corresponds to an option that you can choose. The decision you must make is to choose one option. The other participant will also have to choose, independently of you, among their options, which correspond to the columns of the table. That is, you choose among rows, while the other participant chooses among columns. However, to simplify things, the experiment is programmed in such a way that all the participants - including the person with whom you are matched - see their decision as shown in the example. That is, each of you will be presented with your possible actions in the rows of the table, and your experimental points will be shown in red. At the time of choosing, you will not know the option chosen by the other participant, and when the other participant is choosing their option, they will not know the option that you have chosen.

The number of experimental points you earn in each of the rounds depends on the option you have chosen and the option that the other participant has chosen.

The table of experimental points you see below is an example of what you will see in each of the rounds.

Example:



For example, if this round is chosen at random and you select the first option (row) and the other participant selects the second option (column), you will obtain 20 experimental points, and the other participant will receive 12 experimental points.

As another example, if this round is chosen at random and you select the third option (row) and the other participant selects the first option (column), you will obtain 18 experimental points and the other participant will receive 14 experimental points.

These are just two examples to better understand how decisions affect the experimental points you can earn and do not intend to suggest what decisions you should make.

To make a selection, click on the white button next to the desired option. Then, the button will turn red to indicate which option you have selected. Once you have chosen an option, the choice is not final, and you can change your selection as many times as you want by clicking on another button until you press the "OK" button that will appear in the lower right corner of each screen. Once you have clicked "OK", the selection will be final, and you will proceed to the next round. You will not be able to move to the next round until you have chosen an option and clicked "OK". You will not have any time restrictions. Take as much time as you need in each round. When all of you have made your decisions in each of the 14 rounds, we will explain the second part of the experiment.

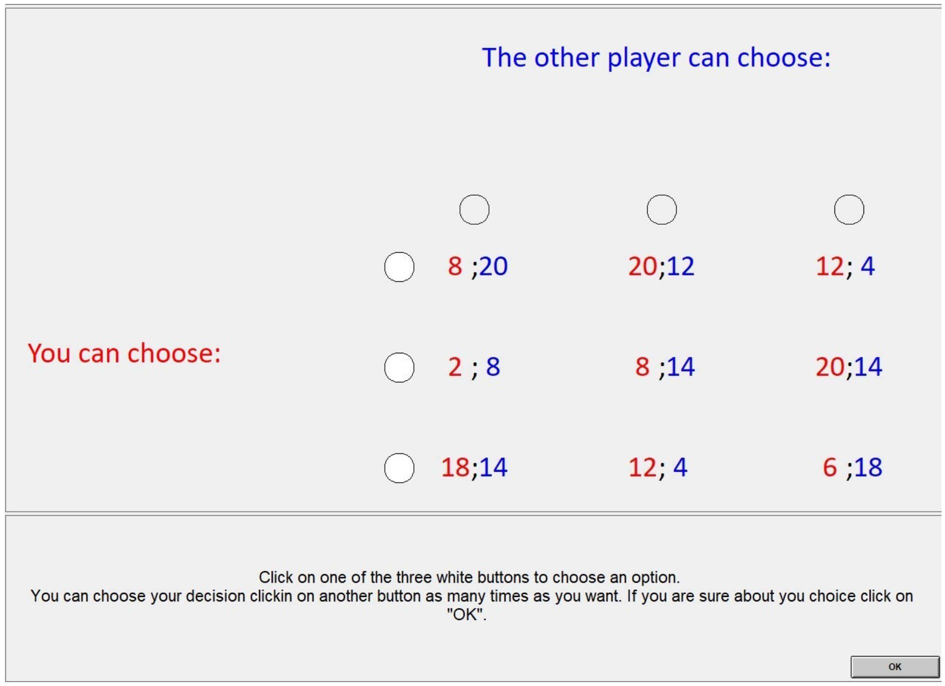
Summary:

- Your experimental points will be shown in red, and the experimental points of the other participant will be shown in blue.

- You will participate in 14 different rounds. In each of the rounds, the table of experimental points will be different, and you will be paired with a different participant chosen at random from this session.

- In each round, you can choose among three different options (rows), and the experimental points that you earn depend on the option you select, the option that the other participant selects, and whether that round is chosen at random at the end of the experiment.

We will start the first part of the experiment in a few moments. Before starting the first part, you will see a new example, and you will have to answer several questions. If you have any questions or need help at any time during the experiment, please raise your hand, and one of the investigators will talk to you.



Questions:

1.Please write the experimental points you would earn in this round, in case this round is randomly chosen for payment, if you chose the second option and the other participant chose the third option.

2.Please write the experimental points the other participant would earn in this round, in case this round is randomly chosen for payment, if you chose the third option and the other participant chose the second option.

3.Please say if it is false or true the following statement: “Two rounds will be randomly selected for payment. The two rounds can be from part 1, from part 2, or 1 from part 1 and the other from part 2.”

SECOND PART OF THE EXPERIMENT:

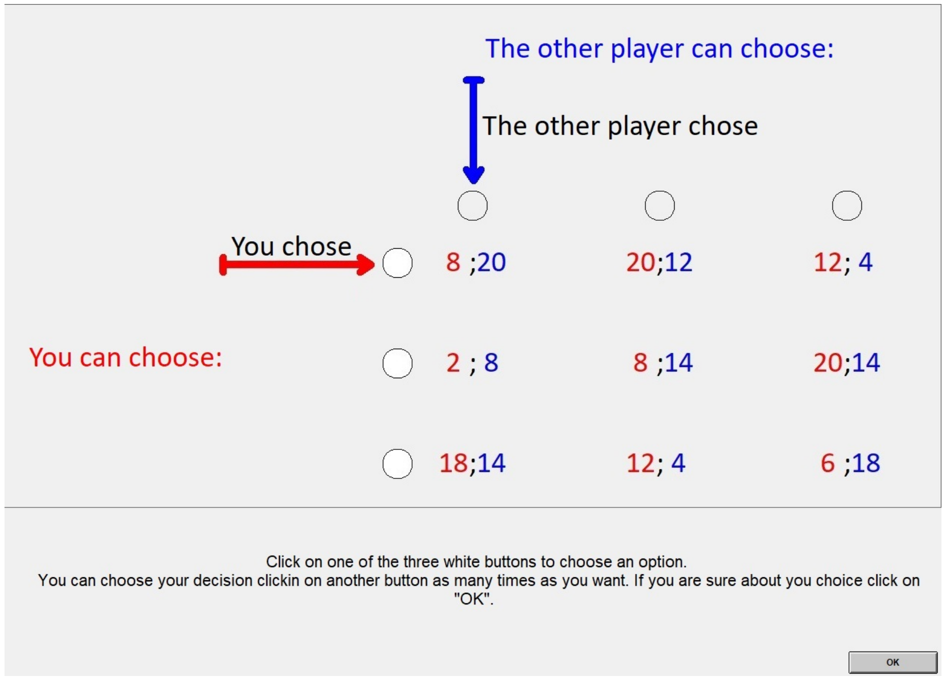
The second part of the experiment also consists of 14 rounds and will work in a similar way as the first part. That is, the tables of experimental points that you will see in each of the 14 rounds in this second part will be the same as those you saw in the first part of the experiment. As in the first part, in each of the 14 rounds, you will be paired with a participant chosen at random from this session. However, in each of the rounds, the other participant with whom you have been paired in this part does not have to be the same as the participant with whom you were paired in the first part. The pairing is performed again at random. In each of the rounds, the other participant, chosen at random, will be different, so you will never be paired with the same participant more than once.

As in the first part, both you and the other participant can choose among three possible options. The experimental points that you can earn in each of the rounds depend on the option that you select and the option that the other participant selects, as well as on whether that particular round is chosen at random at the end of the experiment.

Unlike the first part, when you see the table of experimental points, you can also observe the option that you chose in the first part and the option that was chosen in the first part by the participant with whom you are paired in this part. The option that you both chose in the first part will be indicated by an arrow and will say "You chose" and "The other chose". The information you observe will be the same for the participants with whom you are paired.

The table of experimental points you see is an example of what you will see in each of the rounds.

Example:



As in the first part, for example, if this round is chosen at random and you select the first option (row) and the other participant selects the second option (column), you will earn 20 experimental points, and the other participant will earn 12 experimental points.

As another example, if this round is chosen at random and you select the third option (row) and the other participant selects the first option (column), you will obtain 18 experimental points, and the other participant will receive 14 experimental points.

These are just two examples to better understand how decisions affect the experimental points you earn and are not intended to suggest what decisions you should make.

Unlike the first part, in this part of the experiment, you can observe, as indicated in the example, which option you chose and which option the other participant chose in the first part. For example, in the example table, you chose the second option (row), and the other participant chose the second option (column). The other participant can also observe the option you chose and the option he/she chose; you both have the same information. Now you will have to make a choice again.

You can make your decision in the same way as in the first part, by clicking on the button of the option you want to choose and confirming by pressing "OK". You will not have any time restrictions. Take as much time as you need in each of the rounds.

When all of you have made your decisions in each of the 14 rounds, the experiment will end.

Summary:

- Your experimental points will be shown in red, and the experimental points of the other participant will be shown in blue.
- You will participate in 14 different rounds. In each round, the table of experimental points will be different, and you will be paired with a different participant chosen at random from this session.
- Unlike the first part, you can now see which option you chose in the first part and which option the other participant chose in the first part. The other participant will also be able to observe the option that he/she chose, as well as the option that you chose.
- In each round, you can choose among three different options (rows), and the experimental points depend on the option you have chosen, the option chosen by the other participant, and whether that round is chosen at random at the end of the experiment.

We will start the second part of the experiment in a few moments. If you have any questions or need help at any time during the experiment, please raise your hand, and one of the investigators will talk to you.