

Alós-Ferrer, Carlos; Ritschel, Alexander

Working Paper

Multiple behavioral rules in Cournot oligopolies

Working Paper, No. 331

Provided in Cooperation with:

Department of Economics, University of Zurich

Suggested Citation: Alós-Ferrer, Carlos; Ritschel, Alexander (2019) : Multiple behavioral rules in Cournot oligopolies, Working Paper, No. 331, University of Zurich, Department of Economics, Zurich

This Version is available at:

<https://hdl.handle.net/10419/202426>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



**University of
Zurich**^{UZH}

University of Zurich
Department of Economics

Working Paper Series
ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 331

Multiple Behavioral Rules in Cournot Oligopolies

Carlos Alós-Ferrer and Alexander Ritschel

July 2019

Multiple Behavioral Rules in Cournot Oligopolies*

Carlos Alós-Ferrer[†]
University of Zurich

Alexander Ritschel
University of Zurich

This Version: July 2019

Abstract

We show that economic decisions in strategic settings are co-determined by multiple behavioral rules. A simple model of intra-individual behavioral heterogeneity predicts testable differences depending on whether rules share a common prescription (alignment) or not (conflict), a classification which is ex ante observable. The predictions include non-trivial response time interactions reflecting the nature of the underlying processes, hence the model is not an as if explanation. In a laboratory experiment and two replications on Cournot oligopolies, we find direct evidence showing that decisions arise from the interaction between a deliberative myopic best reply rule and a more intuitive imitative rule.

JEL Classification: C72 · C92 · D03

Keywords: Multiple behavioral rules · Cournot oligopoly · Best Reply · Imitation · Reinforcement

Working Paper. This is an author-generated version of a research manuscript which is circulated exclusively for the purpose of facilitating scientific discussion. All rights reserved. The final version of the article might differ from this one.

*Alexander Ritschel was financed by the Research Unit “Psychoeconomics” (FOR 1882) of the German Research Foundation (DFG).

[†]Corresponding author: carlos.alos-ferrer@econ.uzh.ch. Department of Economics, University of Zurich. Blümlisalpstrasse 10, 8006 Zurich, Switzerland.

1 Introduction

How are economic decisions made? This question becomes especially relevant once one moves away from the assumption of full rationality, or the proximate assumption of perfect maximization of stable preferences. Behavioral economics has documented a large number of behavioral inconsistencies that cannot be accommodated by neoclassical models (e.g., Kahneman, 2003). The microeconomics literature contains extensive and mature strands analyzing boundedly rational behavior in order to explain such inconsistencies. For instance, models of learning and evolution have analyzed different behavioral rules capturing specific deviations from rationality, as e.g. imitation, satisficing behavior, or reinforcement learning (e.g. Kandori et al., 1993; Vega-Redondo, 1997; Erev and Roth, 1998). To date, however, these behavioral rules have been treated as “black boxes” and mostly studied in isolation, while the neoclassical literature has kept studying fully rational, maximizing agents.

Obviously, economic agents do try to maximize certain objectives. Even more obviously, they frequently fail to do so and follow different impulses instead. In other words, the truth might be in the middle, and it might be worth to take both views into account when studying human behavior. In particular, we suggest that economic decisions might be often best viewed as the result of the interaction of multiple decision rules, painting a picture of *intra-individual heterogeneity*.

In this work, we focus on one of the most prominent settings capturing strategic interactions in markets, the Cournot oligopoly. Previous evidence, both theoretical and experimental, suggests that two particular behavioral rules are important in this context. On the one hand, *myopic best reply* captures one-step payoff maximization and acts as a first-order proxy of deliberative thinking. On the other hand, *imitation* of successful behavior is the natural candidate as an alternative rule governing behavior. Vega-Redondo (1997) showed that if firms follow imitative behavioral rules and make infrequent mistakes, the system converges to the Walrasian equilibrium (in the sense of stochastic stability). The reason is that imitation in Cournot oligopolies mimics maximization of relative payoffs (Schaffer, 1989), which can destabilize the Cournot-Nash equilibrium but not the Walrasian one. Alós-Ferrer and Ania (2005) showed that this result extends beyond Cournot oligopolies to a wide class of economic interactions (aggregative games). These results have been shown to be empirically relevant in the behavioral laboratory. A number of Cournot-oligopoly experiments (Huck et al., 1999; Offerman et al., 2002; Apesteguía et al., 2007, 2010) have found partial convergence to Walrasian outcomes, which has then been interpreted as indirect evidence for the presence of imitative behavior.

Our approach is more direct in that we study individual decisions in themselves, as opposed to long-run convergence. Following the evidence mentioned above, we postulate the existence of (at least) two decision rules, namely imitation of observed, successful behavior and myopic best reply, that is, payoff maximization taking current information

on other players’ behavior as given. In our context it is reasonable to assume that imitation is a more heuristic rule, where individuals react to a more successful action and respond by imitating this action, while myopic best reply is a more deliberative rule which involves active maximization after considering available information.

To examine the multiple behavioral rules we rely on a simple formal model which encompasses and extends a “dual-process diffusion model” previously used to study process multiplicity in individual, non-strategic, binary decisions (Achtziger and Alós-Ferrer, 2014; Alós-Ferrer, 2018). To derive testable predictions, the model borrows ideas from *dual-process theories*, which postulate that the human mind is mainly influenced by two kinds of processes, called *automatic* and *controlled* (see Kahneman, 2003; Alós-Ferrer and Strack, 2014; see also Evans, 2008 and Weber and Johnson, 2009 for detailed reviews). Automatic processes are fast, unconscious, and require few cognitive resources. They capture impulsive reactions and behavior along the lines of stimulus-response schemes. Controlled processes are slow, consume cognitive resources, and are reflected upon (partly) consciously. Although there is a clear analogy between these theories and the duality between full and bounded rationality in economics, the key difference is that dual-process models assume heterogeneity within the individual. For our purposes, the key observation is that imitation, as a boundedly-rational behavioral rule, can be expected to be more automatic than rules assuming explicit payoff maximization as myopic best reply.¹

Our predictions focus on one of the most basic measures of process data, *response times*. Those are a standard tool in psychology and are now slowly being incorporated into the economist’s toolbox (Moffatt, 2005; Rubinstein, 2007, 2016; Achtziger and Alós-Ferrer, 2014; Alós-Ferrer et al., 2016; Alós-Ferrer and Ritschel, 2018; Spiliopoulos and Ortmann, 2018). The key insight allowing for testable predictions is that automatic processes are faster than controlled ones, and hence response times can be used as a direct source of evidence for the involvement of different decision processes. This does *not*, however, mean that one can simply classify decisions in fast and slow according to some exogenous criterion and conclude that one kind of decisions is more automatic. This would be an example of the “reverse inference” fallacy (Krajbich et al., 2015). The problem is that processes, and behavioral rules, are not directly observable—only choices are. Hence, when observing a choice and its associated response time, we cannot know which process has generated them. Each process will result in a distribution of response times (and choices!). However, by exploiting the concepts of conflict and alignment among behavioral rules (i.e., whether they prescribe the same answer or different ones), our model avoids reverse inference while still allowing for specific, non-trivial predictions (on response times conditional on specific types of choices).

¹The key differences with research in cognitive psychology are, first, that the paradigm we focus on is far more complex than those typically encountered in that literature, and, second, that the behavioral rules we are interested in all involve cognitive aspects (as opposed to purely automatic reactions).

The model delivers four kinds of predictions. First, whenever best reply and imitation are in conflict (make different prescriptions), choices where a best reply is selected are slower than when both rules are aligned (make the same prescription). Intuitively, this is because in case of conflict best replies come almost exclusively from the slower best reply rule, while in case of alignment the faster imitation rule also contributes a significant proportion of best replies. Second, in case of conflict, best replies are slower than imitative decisions, essentially because many of the latter arise from the faster imitation rule. Third, in contrast to the case of conflict (and somewhat counterintuitively), in case of alignment best replies are *faster* than other responses. This is because, in this case, the faster imitation rule contributes a large number of apparent best replies. Fourth, there are less best replies in case of conflict than in case of alignment. Also, there are less imitative choices in case of conflict than in case of alignment. This is simply because in case of alignment both behavioral rules favor a common prescription.

We conduct one main laboratory experiment and further analyze data from two replications. The main experiment finds clear evidence for the predictions explained above. We also examine a related experiment using a similar design where each of two treatments can be seen as a (smaller) replication; the treatments differed in a cognitive load manipulation, whose effects are examined in Achtziger et al. (2019). For each of the treatments in that experiment, all prior predictions apply, hence we replicate the analysis and confirm the conclusions of our main experiment. In summary, our results suggest that multiple behavioral rules codetermine behavior in a complex economic setting (Cournot oligopoly), with imitation of past success and myopic payoff maximization being the two main drivers of decisions. This multiplicity occurs at the individual level, that is, behavioral heterogeneity starts within each single decision maker.

Although we have concentrated on myopic best reply and imitation as the two main determinants of behavior, our design allows us to consider other behavioral rules. The first is *positive reinforcement*, i.e. the tendency to repeat successful actions (closely linked to the focus on past performance). The second is *inertia*, which simply means the tendency to repeat the previous action regardless of the previous result. The results suggest that the former (but not necessarily the latter) also plays a role, strengthening the case for multiplicity of behavioral rules.

The remainder of the paper is structured as follows. Section 2 presents a simple formal model and derives our predictions. Section 3 presents the design and the results of our main experiment. Section 4 does the same for the replications. Section 5 discusses the additional behavioral rules given by reinforcement and inertia. Section 6 concludes. Proofs, experimental instructions, and screenshots are in the Appendix.

2 Predictions for Multiple Behavioral Rules

2.1 A Simple Formal Model

Consider a given decision problem, where a decision maker has received some information on the available alternatives. On the basis of possibly-different parts of that information, different behavioral rules deliver prescriptions. Suppose further that only finitely many options are available (as will be the case in the experiments). Denote by X the finite set of options, with typical element $x \in X$.

In the context of our experiments, the information corresponds to the quantities produced and profits earned by all involved firms. This, together with the structure of the game, allows to compute both the action which maximizes payoffs given the actions of other players (myopic best reply) and to observe the action which has led to the largest payoffs in the last interaction (imitative choice). Let x^B denote the myopic best reply and x^I the imitative choice. Assume for simplicity that there are no ties (as will be the case in the experiments). We will assume below that the myopic best reply favors x^B above other options, and that the imitation rule favors x^I above other options. We speak of *conflict* if $x^B \neq x^I$, that is, imitating the best observed payoffs would not result in a best reply, and we speak of *alignment* if $x^B = x^I$.

The following model generalizes the model of Achtziger and Alós-Ferrer (2014), which was restricted to binary choice, to the multiple-alternative case, and further extends the results there and in Alós-Ferrer (2018). The model assumes that two behavioral rules codetermine behavior, a more deliberative one and a more intuitive/impulsive one. For the purposes of the present manuscript, we concentrate on myopic best reply and imitation, but the analysis in this section applies to any two given behavioral rules (see Section 2.3 below).

Let BR and Im denote the myopic best reply and imitation rules, respectively. Which of the two rules will actually determine behavior is a stochastic event. Let $\Delta > 0$ be the probability that the actual response is selected according to imitation, and $1 - \Delta$ the probability that it is selected according to myopic best reply. However, we assume that all rules are stochastic in nature, i.e., they carry an amount of noise, resulting in errors (deviations from the rule's prescription). Note that, hence, myopic best reply can select x^I and imitation can select x^B even in case of conflict, and any of them could select actions $x \neq x^B, x^I$. That is, in case of alignment ($x^B = x^I$) both behavioral rules tend to make the same prescription and in case of conflict ($x^B \neq x^I$) they would make different prescriptions in the absence of noise, but due to behavioral noise they might actually select either option in either case, or a third, different one.

Denote by P^{BR} the probability with which the myopic best reply rule indeed selects the best reply x^B , and by P^{Im} the probability with which the imitation rule selects the alternative with the highest observed payoff, x^I . That is, if $P^{BR}(x)$ and $P^{Im}(x)$ denote the probabilities with which each rule selects $x \in X$, conditional on the rule being the

one which actually determines the response, then $P^{BR} = P^{BR}(x^B)$ and $P^{Im} = P^{Im}(x^I)$. Our first assumption is as follows.

(P1) For each decision situation,

$$P^{BR} > P^{BR}(x) \forall x \in X, x \neq x^B \quad \text{and} \quad P^{Im} > P^{Im}(x) \forall x \in X, x \neq x^I.$$

This is a minimal consistency condition which simply declares that the prescription of a rule is indeed the rule's most frequent selection, but it is a rather mild one, since for the multi-alternative case it does not even imply that the prescription is selected more than half of the time.

Response times are also assumed to be stochastic. Let $R^B = E[RT|BR]$ and $R^I = E[RT|Im]$ denote the *expected* response times conditional on the response being selected by the myopic best reply or the imitation rule, respectively. For simplicity, we assume that expected response times do not depend on the actually-selected response. Naturally, since imitation is thought to be more automatic, hence faster *in expected terms*, we assume

(R) $R^B > R^I$.

For some of the results below, we will further assume that

(P2) $P^{Im} > P^{BR}$,

i.e. the deliberation process behind best reply (computing the myopically optimal behavior) is noisier than the stimulus-response process behind imitation (copying the action with the largest observed payoff), while the latter is more *consistent*. This is natural since imitation is assumed to be more automatic (closer to a stimulus-response process).

A simple way to think of the model is to conceive of the imitation rule as a swift cognitive shortcut, which selects the action with the largest observed payoff quickly and very frequently, while the myopic best reply rule is a slow, deliberative process which depends on actual computations and is hence less consistent.

For the binary-choice case, the model in Achtziger and Alós-Ferrer (2014) has been given a micro-foundation in Alós-Ferrer (2018) as the *dual-process diffusion model* or DPDM. In this model, the processes are instantiated as diffusion processes as in the drift-diffusion model (DDM) of Ratcliff (1978) and Ratcliff and Rouder (1998), which has been recently further analyzed by Fudenberg et al. (2018) and is standard in cognitive psychology and neuroscience (e.g. Shadlen and Shohamy, 2016). In this model, evidence accumulation (internal to the decision maker) is captured as a diffusion process with a trend μ and two barriers. Whether the process chooses an option or the other corresponds to whether the upper or the lower barrier is hit first. The response time is the time at which the first barrier is hit. Alós-Ferrer (2018) shows that, in the DPDM, assumptions (P1), (P2), and (R) are implied if one simply assumes that the drift rate of the more automatic process is larger in absolute value than the drift rate of the more deliberative process, capturing that the former is swifter than the latter.

2.2 Predictions

Since all our formal results translate directly into experimental hypotheses, we label them accordingly for convenience (H1, H2, etc). The first testable prediction of the model concerns the comparison of conflict and alignment. Recall that, by committing *ex ante* to which behavioral rules are of interest, we can identify situations of conflict and alignment before data collection. The first prediction states that the response time of best replies must be strictly larger in situations of conflict than in situations of alignment. Since this prediction arises exclusively from process multiplicity, it essentially constitutes a “smoking gun” test on the presence of multiple processes.

Theorem 1. *Under (P1) and (R),*

(H1) the expected response time of best replies in case of conflict is strictly longer than the expected response time of best replies in case of alignment.

The intuition for Theorem 1 is as follows (all proofs are in the online appendix). Independently of whether the decision problem corresponds to conflict or alignment, the best reply rule delivers the same proportion of best replies, which are relatively slow. In case of conflict, the imitation rule favors the imitative choice, which is not a best reply, and hence typically contributes relatively fewer (fast) best replies. In case of alignment, the imitation rule actually favors the best reply, and hence typically contributes relatively many (fast) best replies. Hence, one obtains faster best replies under alignment than under conflict.

The model also makes more nuanced predictions for the response times of best replies and other responses. Those amount to a non-trivial interaction between responses (best replies, imitative choices, or other alternatives) and cognitive situations (conflict or alignment). Specifically, best replies must be slower on average than imitative choices in case of conflict, but in case of alignment (where best replies are also imitative choices), they must be faster than other choices. This parallels the prediction of Achtziger and Alós-Ferrer (2014) and Alós-Ferrer (2018) that in situations with normatively correct answers errors are fast in case of conflict but slow in case of alignment. This asymmetry goes beyond simple informal statements that intuitive responses should be faster, which might hide a reverse inference fallacy (Krajbich et al., 2015), and serves as a test of the basic structure of the model. The next result gathers the predictions.

Theorem 2. *Assume (R).*

(H2) Under (P1), in case of conflict, the expected response time of best replies is larger than the expected response time of imitative choices (choosing the alternative with highest observed payoff).

(H3) Under (P2), in case of alignment, the expected response time of best replies is shorter than the expected response time of other choices.

The intuition behind Theorem 2 is as follows. The (slow) best reply rule favors the best reply alternative and the (fast) imitation rule favors the imitative choice. Those two alternatives are different in case of conflict, and hence best replies end up being on average slower in this case. In case of alignment, the two alternatives coincide but by (P2) the fast imitative process contributes more of them than the best reply rule, hence in expected terms best replies end up being on average faster. In other words, in case of alignment, the imitation rule acts as a quick and efficient shortcut to identify the best reply while the more error-prone best reply rule contributes relatively more (slow) non-best-reply answers. Hence, conditional on a best reply not being observed, it is more likely that the response is generated by the slower best reply rule.

Last, the model also makes predictions for the proportion of best replies and imitative choices comparing the cases of conflict and alignment, which we summarize in the following result.

Theorem 3. *Under (P1),*

(H4a) the proportion of best replies is strictly smaller in case of conflict than in case of alignment, and

(H4b) the proportion of imitative choices is strictly smaller in case of conflict than in case of alignment (when they are also best replies).

The intuition for Theorem 3 is immediate. In case of alignment, both behavioral rules favor the same option, in the sense of being the one selected most often. That option is simultaneously a best reply and an imitative choice. In case of conflict, the myopic best reply rule still favors best replies, but the imitation rule now favors a different option, which is imitative but not a best reply. Even though each rule might still select the option favored by the other rule in case of conflict, it does so less often. Hence, in case of alignment the common prescription obtains more often than any of the individual choices in case of conflict.

2.3 Beyond Best Reply and Imitation

The model applies to any situation where the researcher can reliably identify two behavioral rules (or decision processes in the sense of psychology) as the main determinants of decisions. It extends beyond the case of best reply and imitation, although we have formulated it in those terms here for concreteness. For instance, the model encompasses the analysis and experimental results in Achtziger and Alós-Ferrer (2014), which considered Bayesian belief updating vs. “win-stay, lose-shift” reinforcement learning as in Charness and Levin (2005), or the example in Alós-Ferrer (2018), which considered following external advice vs. the recognition heuristic. Another example is Spiliopoulos (2018), who used the model in Alós-Ferrer (2018) to study win-stay, lose-shift vs. more sophisticated (cognitive) heuristics in a repeated game played against computer algorithms.

The key assumption to apply the model is that one of the postulated rules or processes can be identified as more deliberative and the other as more intuitive, where the formal meaning of these labels is as given by (R) and (P2). Assumption (P1) follows simply from the interpretation of a behavioral rule as a stochastic mapping favoring a particular type of prescription. However, we remark that, given (P1), predictions (H1) and (H2) only depend on (R), that is, they hold even if (P2) does not. Likewise, predictions (H4a) and (H4b) only depend on (P1). Prediction (H3) is the only one requiring (P2) (but, interestingly, not (P1)).²

As an illustration, suppose one identifies the more deliberative rule (or more reflective process) with more rational behavior, as would be the case, e.g., if this rule was stated as a summary of the normative predictions in the decision problem at hand. We would then be justified in calling the action prescribed by this rule “correct” and any other action an “error.” In particular, correct responses would play the role of myopic best replies above. Suppose the alternative behavioral rule reflects a heuristic or bias, which the modeler hypothesizes to interact with normative behavior; in psychological terms, due to (R) this rule should be viewed as a more automatic process. Then, in case of conflict the imitative choices above correspond to a particular type of errors, namely to those following the heuristic’s prescription: call them “heuristic errors.” There are other errors, however, which are not of the heuristic type. In terms of the model, those are due to behavioral noise. In this setting, the more automatic behavioral rule is actually a quick, efficient shortcut to the correct response, but only in case of alignment. In case of conflict, it captures a quick path to a particular type of error.

With this interpretation, the predictions above translate as follows. First (H1), response times *of correct responses* should be longer (on average) in case of conflict than in case of alignment. Second (H2), conditional on conflict situations, heuristic errors should be faster (on average) than correct responses. Third (H3), conditional on alignment situations, errors should be slower (on average) than correct responses. Note that the two latter predictions are of different nature, since (H2) refers only to errors of a particular type (the ones corresponding to the heuristic or automatic process’ predictions), while (H3) refers to all errors in case of alignment (where, actually, there are no heuristic errors). Last (H4a), there should be more correct responses in case of alignment than correct responses in case of conflict, and (H4b), there should be more correct responses in case of alignment than heuristic errors in case of conflict (but not necessarily more than errors of all types taken together).

²In particular, the results obtained here rest on weaker assumptions than those in Achtziger and Alós-Ferrer (2014) or Alós-Ferrer (2018), in addition to the fact that the model at hand allows for more than two alternatives.

2.4 Model Extension: Non-Decision Time and Conflict-Dependent Process Selection

Theorem 1 predicts a slow-down of certain responses under conflict compared to alignment. It is worth noticing that this prediction corresponds to the well-known “Stroop Effect” discussed in psychology (Stroop, 1935; MacCleod, 1991), which describes a slow-down of (correct) responses when one is asked to name the color that a word is printed in but that word happens to name a different color (e.g., “Red” printed in blue) compared to when the word names the color it is printed in (e.g., the word “Red” printed in red). However, work in psychology typically assumes that this and similar response-times effects are due to central executive functions of the brain related to the detection and resolution of conflict among elementary responses, which tax cognitive resources and require time (Bargh, 1989; Baddeley, 1992; Baddeley et al., 2001), but enable the inhibition of automatic responses in case of conflict. These functions have been linked to early activity in the Anterior Cingulate Cortex (see, e.g., Nieuwenhuis et al., 2003; De Neys et al., 2008; Achziger et al., 2014).

It is worth noticing that our model does not assume such a difference in response times and Theorem 1 holds in its absence, providing an alternative (or complementary) explanation for the Stroop effect in psychology. It is, however, easy to extend the model to account for the additional insights described above. Let $i \in \{A, C\}$ denote alignment or conflict, respectively, and add a “non-decision time” t_i to the response time which depends on conflict vs. alignment and is such that $t_C \geq t_A$. At the same time, since conflict detection enables the inhibition of automatic responses, an extended model should distinguish the probability of the latter depending on conflict or alignment, i.e. replace Δ with Δ_i while assuming $\Delta_C \leq \Delta_A$. It is easy to see that all our results hold in the extended model.

Theorem 4. *Consider the extended model and assume (P1), (P2), (R), $t_C \geq t_A$, and $\Delta_C \leq \Delta_A$. Then (H1), (H2), (H3), (H4a), and (H4b) hold.*

The extension, however, disciplines the model in sensible ways. For instance, an analogous proof to that of Theorem 1 shows that the expected response time of imitative choices in case of conflict is strictly shorter than the expected response time of best replies (which are also imitative choices) in case of alignment. However, this prediction does not necessarily hold in the extended model, since non-decision times are longer in case of conflict and hence the comparison of total response times would be undetermined.

3 Main Experiment

3.1 Experimental Design and Procedures

In our main experiment, participants interacted in 4-player Cournot oligopolies (tetrapolies). We conducted four sessions with 32 participants each for a total of $N = 128$ (82 females;

median age 22 years) at the Cologne Laboratory for Economic Research (CLER). The experiment was programmed with z-Tree (Fischbacher, 2007) and participants were recruited using ORSEE (Greiner, 2015). We excluded students majoring in economics, psychology, and business, as they might have been taught game-theoretic concepts which might influence their behavior. A session lasted around 90 minutes and average earnings were 13.59 EUR, including a show-up fee of 2.50 EUR.

Each participant competed in three different Cournot oligopolies (parts), which lasted for 17 periods each. Initially, players were matched in groups of four to play the first tetrapoly (Part 1). After 17 periods, players were rematched in new groups of four and the oligopoly payoffs (demand function) were changed (Part 2). After 17 further periods, players were rematched again and played a third oligopoly with new payoffs (Part 3). To increase the number of fully-independent observations, rematching was done within 16 pre-determined blocks of 8 participants each. Identities within a part were always anonymous and could not be traced back to previous parts. In each new part, at least two players in the group were different from the previous group. The sequence of the different oligopolies was varied across sessions.

Each oligopoly was implemented through a payoff table derived from a linear inverse demand function of the form $P(Q) = a - Q$, where P is the price, a the saturated demand, Q the total quantity in the market, and linear costs are normalized to zero. A neutral framing was used and neither firms nor quantities were mentioned. We reduced the action space to four possible actions (A , B , C , and D). Hence, the game is given by a $4 \times 4 \times 4 \times 4$ payoff table, which by symmetry can be reduced to a 4×20 table, with four rows for the possible actions and 20 columns (labeled AAA to DDD) for the opponents' actions (independently of their identity). Payoffs were expressed in points (rounded to the nearest integer), with an exchange rate of 18 Eurocents per 1000 points. The points achieved in all 51 rounds were accumulated and paid at the end of the experiment (all decisions were paid). The payoff table was permanently visible in the upper part of the screen during the corresponding part of the experiment. Example screenshots and instructions are presented in the (Online) Appendix.

In order to focus on the interaction between myopic best reply and imitation, we highlighted the information required to implement both rules. Myopic best reply implements maximization within the column corresponding to the actual actions of the opponents in the previous period. For all rounds except the first one within each part, that column was highlighted. Thus, determining a myopic best reply required comparing four numbers only. For each round except the first, participants were also given feedback on the actions and profits of the group members in the previous period, making imitation feasible. As a robustness control, to make sure that presentation effects were minimized, we included two treatments which differed only on how that information was presented. In Treatment *FullInfo*, the choice and points of all other group members were presented in separate boxes, in addition to a box displaying the own choice and received points, and the box with the highest point amount was highlighted. In Treatment *BestOnly* only the

own choice and points plus an additional (highlighted) box were shown, with the latter displaying the choice with the largest amount of points in the previous round (and the corresponding points). Note that in the FullInfo treatment both imitation and myopic best reply involve comparing four numerical quantities, making the mechanical aspects of the rules as comparable as possible. In contrast, the BestOnly treatment closely reproduces the idea of “imitate the best” as described, e.g., by Vega-Redondo (1997), i.e. choosing the action with the highest profit in the previous round. The treatments were implemented between subjects, with half the subjects in each treatment in every session. As we will see below, results are not affected by the differences in information presentation.

The rationale for the experimental implementation is as follows. First, we discretized the action space to make the postulated behavioral rules (myopic best reply and imitation) both feasible and comparable. A continuous- or large-action space would have turned myopic best reply into an abstract maximization problem, while imitation would remain a discrete, intuitive rule. By choosing a discrete setup we go against our hypotheses and reduce the conceptual distance between the two behavioral rules.

Second, in contrast to previous experiments with Cournot oligopolies (e.g., Offerman et al., 2002; Apesteguía et al., 2007), we are interested in behavioral correlates of individual actions, rather than on eventual convergence. If and when convergence occurs, there is no further variance in the behavioral (choice) data, and response times become meaningless as participants mechanically repeat a fixed action. Hence, we were interested in data before convergence occurred. To maximize usable data, we implemented three parts (oligopolies) with rematching of participants, reassignment of identities, and changed payoff tables (computed with different demand functions and different quantities underlying the four actions). Further, the ordering of the quantities (A to D) changed with each part, that is, in some parts the assignment of quantities to letters was increasing and in some it was decreasing.³ The second and third parts always had a different payoff table and a reversed ordering of the quantities with respect to the previous part.

Third, by the same reasoning data would be meaningless if and when collusion occurred. Rematching, working with shorter oligopolies, and changing payoff tables across parts already diminish the likelihood of collusion and increase the variance in behavioral data. Additionally, while previous experiments with Cournot oligopolies have typically focused on triopolies to increase the number of independent observations for a given number of participants, we chose to focus on tetrapolies because larger groups make collusion less likely and ensure higher outcome volatility (see Huck et al., 2004).

³Payoff table 1: $P(Q) = 150 - Q$, $A = 37.5$, $B = 33.25$, $C = 30$, $D = 18.75$ (or reversed); Payoff table 2: $P(Q) = 175 - Q$, $A = 43.75$, $B = 38.875$, $C = 35$, $D = 21.875$ (or reversed); Payoff table 3: $P(Q) = 200 - Q$, $A = 50$, $B = 44.5$, $C = 40$, $D = 25$ (or reversed).

Table 1: Overview of Prescribed Actions

	AAA			AAB			AAC			AAD			ABB			ABC			ABD			ACC			ACD			ADD		
	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In
A	D	A	A	D	A	A	D	A	A	C	A	A	D	A	A	C	A	A	C	A	A	C	A	A	B	A	A	A	A	A
B	D	A	B	D	A	B	D	A	B	C	A	B	D	A	B	C	A	B	C	A	B	C	A	B	B	A	B	A	A	B
C	D	A	C	D	A	C	D	A	C	C	A	C	D	A	C	C	A	C	C	A	C	C	A	C	B	A	C	A	A	C
D	D	A	D	D	A	D	D	A	D	C	A	D	D	A	D	C	A	D	C	A	D	C	A	D	B	A	D	A	A	D

	BBB			BBC			BBD			BCC			BCD			BDD			CCC			CCD			CDD			DDD		
	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In	BR	Im	In
A	C	A	A	C	A	A	B	A	A	C	A	A	B	A	A	A	A	A	C	A	A	A	A	A	A	A	A	A	A	A
B	C	B	B	C	B	B	B	B	B	C	B	B	B	B	B	A	B	B	C	B	B	A	B	B	A	B	B	A	B	B
C	C	B	C	C	B	C	B	B	C	C	B	C	B	B	C	A	B	C	C	C	C	A	C	C	A	C	C	A	C	C
D	C	B	D	C	B	D	B	B	D	C	B	D	B	B	D	A	B	D	C	C	D	A	C	D	A	C	D	A	D	D

Note: Overview of prescribed actions for each behavioral rule depending on last period's outcome. Cell entries describe the action prescribed by myopic best reply (BR), imitation (Im), and inertia (In) when the player previously chose the action given in the row and the opponents chose the actions given in the column. Shaded entries for BR and Im indicate that the two rules are aligned. Shaded entries for In indicate positive reinforcement.

3.2 Classification of Decisions and Strategy of Analysis

The data set of our main experiment consists of $128 \times 48 = 6,144$ observations. The first decision within each part is always excluded since for that period there is no feedback concerning previous actions and the behavioral rules considered make no prescriptions.

Given the previous actions of all four players, the identification of the prescriptions of the different behavioral rules is straightforward. Table 1 displays the prescriptions of myopic best reply and imitation in the experiment, for the case of decreasing assignment of quantities to letters.⁴ Those prescriptions were identical for all three payoff tables. That is, the table shows the prescription of each behavioral rule when a specific combination of one's own choice (row) and the choice of the other players (column) occurred in the previous round. Whenever myopic best reply is in alignment with imitation (that is, both prescribe the same action), the corresponding cells are shaded in gray. Hence, unshaded entries indicate conflict between myopic best reply and imitation.

For reference, Table 1 includes also the prescriptions of *inertia*, that is, simply repeating last period's action independently of payoffs (e.g., Alós-Ferrer et al., 2016). Shaded entries in the inertia column indicate when those prescriptions coincide with those of imitation. This is of particular interest because in this case both rules are equivalent to *positive reinforcement* (imitating yourself if you obtained the largest payoffs). This information will be used in Section 5 below.

Given Table 1, for periods 2–17 within each part, we can classify each actual decision of each participant depending on whether it is consistent with myopic best reply or imitation (or inertia). The Venn diagram in Figure 1 gives a descriptive classification of the actual decisions in the experiment. Actions can be classified as imitation, myopic best reply, or inertia, as belonging to any of the intersections (alignments), or as being inconsistent with all of them (unclassified). As was to be expected, the majority of the

⁴For the analysis of the data, the case of increasing assignment of quantities was simply recoded.

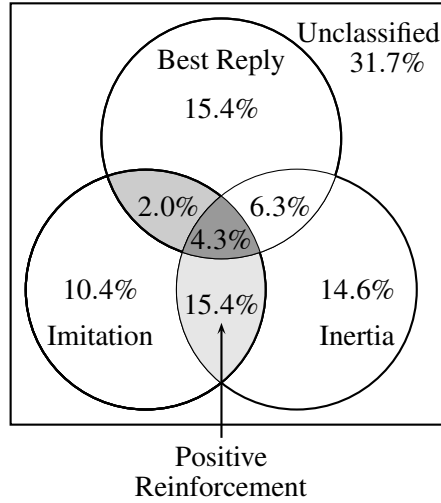


Figure 1: Overview of observations and their classification, Main Experiment.

6,144 decisions were made in conflict situations (5,010; 81.54 percent). Of those, 26.57, 31.64, and 41.80 percent were myopic best replies, imitative decisions, or other choices, respectively. However, there were enough decisions made in case of alignment (1,134; 18.46 percent) to enable a meaningful analysis. Of those, 34.13 percent were myopic best replies (and hence also imitative).

In order to test our hypotheses, we will initially conduct non-parametric tests. For instance, we can test whether decisions compatible with one behavioral rule are faster than those compatible with another decision rule, conditional, e.g., on conflict among the rules (Hypotheses H2, H3). To do so, we look at all situations where the two rules conflict and build two sets of decisions for each individual, those where the prescription of the first rule was followed, and those where the prescription of the second rule was followed. Then we apply the appropriate test (in this case, a Wilcoxon Signed-Rank test).

For the analysis we consider the matching block the appropriate unit of observation, i.e. observations of all subjects who interacted anonymously with each other throughout all 3 parts are pooled into one observation. Since participants were separated into $N = 16$ different blocks (8 in each treatment) and were rematched only within those blocks, this guarantees completely independent observations. For each block, we compute the relative frequencies of choices and the average response times when following a given behavioral rule, conditional on conflict or alignment of myopic best reply and imitation.⁵

⁵A case can be made for individual observations as the appropriate unit of analysis. Following the logic of stochastic evolutionary models (Blume, 1993; Kandori et al., 1993; Vega-Redondo, 1997; Alós-Ferrer and Ania, 2005), behavioral rules have a Markovian structure, i.e. they are mappings from information (outputs and profits in last period) to actions. Under this assumption, how exactly the input of the behavioral rule is generated is irrelevant. Hence, the fact that participants were part of tetrapolies which themselves were subgroups of certain blocks plays no role, for we are testing relative frequencies and response times which are generated *after* observation of the input, and tests condition on the relevant categories of inputs. Our conclusions were unchanged when conducting tests at the individual level (considering only those average response times with at least two observations per individual).

Before proceeding to the main analysis, we comment on the informational treatments. Those served as a robustness check to ensure that mere presentational effects, as salience of the maximum observed payoffs, did not significantly affect response times or drive behavior toward imitation. A block of 8 participants made on average 125.13 imitation decisions in the BestOnly treatment and 121.38 in the FullInfo treatment, which was not significantly different according to a Mann-Whitney-Wilcoxon (MWW) test ($N = 16$, $z = 0.735$, $p = .4622$). The average response time of imitation decisions was 10.38 s in the BestOnly treatment and 10.36 s in the FullInfo treatment (MWW, $N = 16$, $z = 0.210$, $p = .8336$). There were also no differences for myopic best replies. Hence, for the remainder of the analysis we will pool the data of both treatments.

3.3 Results

Figure 2 illustrates all our predictions for the main experiment. Average response times are shown on the left-hand side, and choice frequencies on the right-hand side. We now discuss all predictions as depicted in the figure, reporting the corresponding non-parametric tests (a regression analysis is discussed below). Note that our hypotheses yield specific directional predictions which would allow us to rely on one-sided p -values. However, we will conservatively report two-sided p -values.

Prediction (H1) serves as a first test of the presence of several, distinct behavioral rules. Myopic best replies, the prescription of the more deliberative behavioral rule, should be slower in case of conflict with imitation than in case of alignment. This corresponds to the comparison between the average response times of best replies in conflict and in alignment in Figure 2. The prediction is confirmed by the data: myopic best replies are slower in conflict (mean 12.38 s) than in alignment (mean 10.46 s), with the differences being highly significant according to a Wilcoxon-Signed-Rank (WSR) test ($N = 16$, $z = 2.947$, $p = .0032$).

Predictions (H2) and (H3) constitute a test of the nature of the involved processes and of the dual-process structure of the interaction. Essentially, myopic best replies should be relatively slow in case of conflict but relatively fast in case of alignment. Specifically, (H2) states that myopic best replies are slower than imitation decisions in conflict situations. As predicted, myopic best reply decisions are slower (average 12.38 s) than imitative choices (average 10.36 s) when the processes make different prescriptions, confirming the relatively more automatic nature of imitation decisions (compare the two left-most bars in the left-hand side of Figure 2). The difference is highly significant according to a WSR test ($N = 16$, $z = 3.361$, $p = .0008$). (H3) states that in case of alignment, myopic best replies (which are also imitative in this case) should be faster than other decisions. As predicted, myopic best replies (average 10.46 s) are significantly faster than other decisions (average 13.63 s; WSR, $N = 16$, $z = -3.258$, $p = .0011$).

The remaining two hypotheses concern relative choice frequencies. (H4a) states that myopic best replies should be less frequent under conflict than under alignment (when

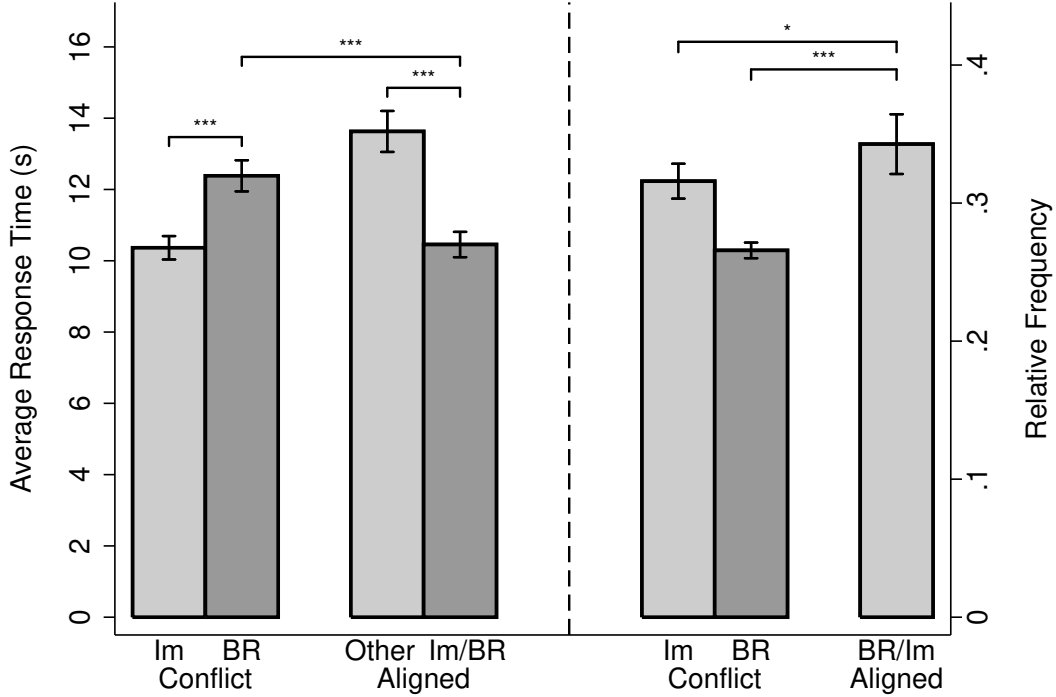


Figure 2: Average response times and choice frequencies, Main Experiment.

Note: Left-hand side: Average response times of imitation, myopic best reply, and other choices conditional on conflict and alignment. Right-hand side: Relative frequency of imitative choices (Im) and myopic best replies (BR) conditional on conflict and alignment. Stars indicate the significance of Wilcoxon Signed-Rank tests. * $p < .1$, ** $p < .05$, *** $p < .01$.

they are also imitative choices). This is illustrated in the right-hand side of Figure 2. In case of conflict, participants chose myopic best replies, on average, 26.57 percent of the time (average of individual averages), compared to 34.27 percent in case of alignment. The difference is highly significant (WSR test, $N = 16$, $z = -2.947$, $p = .0032$). (H4b) states that, in contrast, imitative decisions should be less frequent under conflict than under alignment (when they are also best replies). This is indeed the case, with an average of 31.59 percent of imitative decisions in case of conflict and 34.27 percent in case of alignment, although the difference is not significant with our two-tailed tests (WSR test, $N = 16$, $z = -1.344$, $p = .1788$). We remark, however, that one group successfully colluded during the last part of the experiment. When excluding the corresponding block observation, we observe less imitative decisions in conflict (average 30.97 percent) than in alignment (34.64 percent; two-tailed WSR test, $N = 15$, $z = -1.874$, $p = .0609$). All previous conclusions regarding (H1-H4a) remain unchanged when excluding the block containing the colluding group.

In summary, simple non-parametric tests already confirm our predictions. Hence, our experimental evidence is compatible with the interpretation that multiple behavioral

Table 2: Random Effects Panel Regressions on (log) Response Times, Main Experiment.

ln(ResponseTime)	Model 1	Model 2	Model 3
Conflict	0.1154*** (0.0360)	0.1418*** (0.0327)	0.1431*** (0.0329)
Imitation-Conflict	-0.1809*** (0.0307)	-0.1510*** (0.0255)	-0.1517*** (0.0255)
Other	0.1794*** (0.0424)	0.1735*** (0.0369)	0.1748*** (0.0370)
Other×Conflict	-0.1808*** (0.0406)	-0.1868*** (0.0354)	-0.1879*** (0.0355)
FullInfo Treatment	-0.0210 (0.0559)	-0.0366 (0.0570)	-0.0308 (0.0517)
Collusion		-0.2405*** (0.0525)	-0.2005*** (0.0513)
Constant	2.2418*** (0.0646)	2.6385*** (0.0579)	2.3426*** (0.1495)
Controls	No	Yes	Yes
Demographics	No	No	Yes
R^2	0.0313	0.1327	0.1587

Note: Standard errors, clustered by 16 matching blocks, in parentheses.

rules, i.e. myopic best reply and imitation, codetermine behavior in complex Cournot oligopolies. We view this as a demonstration that complex economic decisions result from the interaction of multiple behavioral rules within individual economic agents.

We now turn to a more detailed regression analysis. Our data forms a perfectly-balanced panel with 48 decisions for each of the 128 participants (total $N = 128 \times 48 = 6,144$). Table 2 reports random effects panel regressions on log-transformed response times.⁶ We (conservatively) cluster standard errors at the block level. Since our hypotheses hinge on the distinction between conflict and alignment, it is important to introduce the appropriate categories in the analysis. The Conflict dummy takes the value 1 when the decision corresponds to a case of conflict between myopic best reply and imitation. To avoid having to rely on *post hoc* tests, we further include the dummy Imitation-Conflict which only considers cases where the imitative choice was selected in conflict situations.⁷ Last, the dummy Other takes the value 1 for choices which are neither imitative nor best replies. Thus, the interaction Other×Conflict indicates choices which are neither imitative nor myopic best replies in case of conflict. Note that the reference group consists of decisions in case of alignment where the myopic best reply (which is also an imitative choice in this case) was selected.

⁶Response times are naturally bounded below by zero and usually present a skewed, non-normal distribution. To account for these features it is common practice to use a logarithmic transformation (Fischbacher et al., 2013; Achtziger and Alós-Ferrer, 2014).

⁷That is, the dummy takes the value 1 for imitative choices in case of conflict, and zero otherwise. Note that, since in case of alignment imitative choices are also best replies, this does not correspond to an interaction in the usual sense of the word.

With this choice of dummies, all our response-times hypotheses can be tested directly in the regressions. Model 1 in Table 2 tests for the basic effects. Models 2 and 3 add further Controls and Demographics⁸ to show that the results are robust. All models include a treatment dummy for the presentation variants, which is never significant. Models 2 and 3 also add a Collusion dummy, taking the value 1 when the subject colluded with other subjects in a Cournot oligopoly. That coefficient is negative and highly significant, showing that the individuals who colluded were, unsurprisingly, fast. The inclusion of that dummy, however, does not affect other results.

(H1) states that best replies should be slower in case of conflict than in case of alignment. The comparison corresponds to the coefficient for Conflict, which is indeed positive and highly significant ($p = .0013$ in Model 1, $p < .0001$ in Models 2 and 3). (H2) predicts that myopic best replies should be slower than imitative choices, a comparison captured by the coefficient for the dummy Imitation-Conflict. The prediction is borne by the data, with the coefficient being negative and highly significant ($p < .0001$ in all models). (H3) predicts that, in case of alignment, best replies should be faster than other responses. The comparison reduces to the coefficient for the Other dummy, which is highly-significant and positive as expected ($p < .0001$ in all models).

The regressions also allow us to examine a number of exploratory questions. The linear combination of the coefficients Other and Other \times Conflict is not significantly different from zero, i.e. in case of conflict we find no differences in response times between myopic best replies and other kinds of non-imitative decisions. In contrast, a linear combination test reveals that imitative decisions are significantly faster than other kinds of non-best-replies in conflict situations ($p < .0001$ in all models). This suggests that this latter category might include choices reflecting higher-level deliberation processes or more complex behavioral rules, as e.g. level- k considerations (best-replying to the anticipated best reply of others, etc; see Alós-Ferrer and Buckenmaier, 2018).

Tables 3 and 4 provide probit panel regressions with myopic best replies and imitative choices as dependent variables, respectively. Standard errors are again clustered at the block level. The independent variables are the Conflict dummy, a treatment dummy, a Collusion dummy, and further Controls and Demographics as in the previous regression models.

Table 3 allows us to parametrically test for Hypothesis (H4a), i.e. the prediction that myopic best replies are less likely under conflict than under alignment. This is confirmed by the negative and highly significant Conflict dummy, which is robust to the addition of Controls and Demographics (Model 1, $p = .0001$; Models 2 and 3, $p = .0002$). Analogously, Table 4 allows us to test for Hypothesis (H4b), i.e. the prediction that imitative choices are also less likely under conflict than under alignment. Although present in the data, this trend is clearly less strong than other predictions. The Conflict dummy is not

⁸Controls consist of a measure for normalized rounds, part 2 and part 3 dummies, and two payoff table dummies for possible medium or high payoffs. Demographics consist of age, gender, and an indicator capturing whether the subjects reported attending a game theory class.

Table 3: Probit Regression Models for Myopic Best Reply, Main Experiment.

Myopic Best Reply	Model 1	Model 2	Model 3
Conflict	−0.2183*** (0.0567)	−0.2163*** (0.0573)	−0.2158*** (0.0574)
FullInfo Treatment	−0.0159 (0.0413)	−0.0258 (0.0409)	−0.0294 (0.0461)
Collusion		−0.1815*** (0.0125)	−0.1573*** (0.0330)
Constant	−0.4092*** (0.0639)	−0.3704*** (0.0620)	−0.4225** (0.1862)
Controls	No	Yes	Yes
Demographics	No	No	Yes
Log Pseudolikelihood	−3612.3172	−3603.9443	−3603.3104

Note: Standard errors, clustered by 16 matching blocks, in parentheses.

Table 4: Probit Regression Models for Imitation, Main Experiment.

Imitation	Model 1	Model 2	Model 3
Conflict	−0.0542 (0.0543)	−0.0754 (0.0470)	−0.0749* (0.0449)
FullInfo Treatment	−0.0159 (0.0752)	0.0282 (0.0802)	0.0120 (0.0739)
Collusion		0.6844*** (0.2523)	0.8105*** (0.1552)
Constant	−0.4557*** (0.0794)	−0.7209*** (0.1323)	−1.2212*** (0.1187)
Controls	No	Yes	Yes
Demographics	No	No	Yes
Log Pseudolikelihood	−3622.6196	−3601.3666	−3596.3697

Note: Standard errors, clustered by 16 matching blocks, in parentheses.

significant in Model 1 ($p = .3184$), which does not control for collusion. The coefficient still misses significance in Model 2 ($p = .1087$) and becomes only weakly significant in Model 3 ($p = .0951$), after adding Controls, Demographics, and the Collusion dummy.

In summary, the regression models confirm our non-parametric analysis while controlling for other features. Taken together, the analyses above provide strong evidence for Hypotheses (H1), (H2), (H3), and (H4a), and weak evidence for Hypothesis (H4b). Hence, we conclude that our main experiment strongly supports our model, suggesting that interacting behavioral rules with qualitatively different properties codetermine behavior in complex economic decisions.

4 Replications

We now present further evidence in the form of a separate experiment. The data is taken from an experiment discussed in Achtziger et al. (2019), which used a closely-related paradigm while manipulating cognitive load between subjects. That is, this experiment

included two separate treatments, each of which can be considered as a replication of our main experiment except for the added cognitive load in one of the two. While Achtziger et al. (2019) discusses the effects of cognitive load across treatments in detail, here we report the analysis of hypotheses (H1)-(H4b) within each treatment.

4.1 Experimental Design and Procedures

The two replications were implemented in the same way as the main experiment, with 3 sessions of 24 participants each yielding $N = 72$ per replication for a total of $N = 144$ (87 females; median age 23 years). The replications were two between-subject treatments, *NoLoad* and *Load*, of an experiment discussed in Achtziger et al. (2019). The *NoLoad* treatment was a replication of the *FullInfo* treatment of the main experiment. The *Load* treatment included an additional cognitive load task but was otherwise identical to the *NoLoad* one. For clarity, we will continue to refer to the first and second replications using the *NoLoad* and *Load* labels, respectively. As in the main experiment participants interacted in 4-player Cournot oligopolies, and there were 51 periods in three blocks of 17 periods each. The cognitive load task consisted of memorizing a seven-digit number (as in Carpenter et al., 2013) before each decision in the Cournot oligopoly. Subjects had 10 s to memorize the number and had to recall the number after the decision. A correct recall was rewarded with 750 additional points.

Apart from the introduction of cognitive load in the *Load* replication, there were only two changes with respect to the main experiment. First, the exchange rate was increased to 20 Eurocents per 1000 points.⁹ Second, subjects were rematched within blocks of 12 participants after each part, which reduces the number of independent observations for our conservative block-level tests.

A session lasted around 85 and 105 minutes in the *NoLoad* and *Load* replication, respectively. Average earnings, including the show-up fee of 2.50 EUR, were 13.61 EUR and 14.06 EUR (excluding the earnings from the cognitive load task).¹⁰ An MWW test at the block level ($N = 12$) shows that average earnings were not significantly different across the two replications ($z = -1.121$, $p = .2623$).

Analogously to the main experiment, Table 1 allows the classification of observations in terms of the prescriptions (favored options) of the behavioral rules of myopic best reply, imitation, and inertia. Figure 3 displays a descriptive overview of all observations by replication and classification according to the behavioral rules. The replications data sets contain 3,456 decisions each. In *NoLoad*, 2,864 (82.87 percent) were under conflict. Of those, 24.02 percent were myopic best replies and 34.78 percent were imitative. The remaining (17.13 percent) were under alignment, of which 43.41 percent were myopic best replies (hence also imitative). In *Load*, 2,892 (83.68 percent) were under conflict.

⁹The exchange rate was increased because the average payoff in the main experiment was slightly below the wage rate demanded by the experimental lab.

¹⁰The average earnings of participants in the *Load* sessions including the payoff from the cognitive load task were 20.12 EUR.

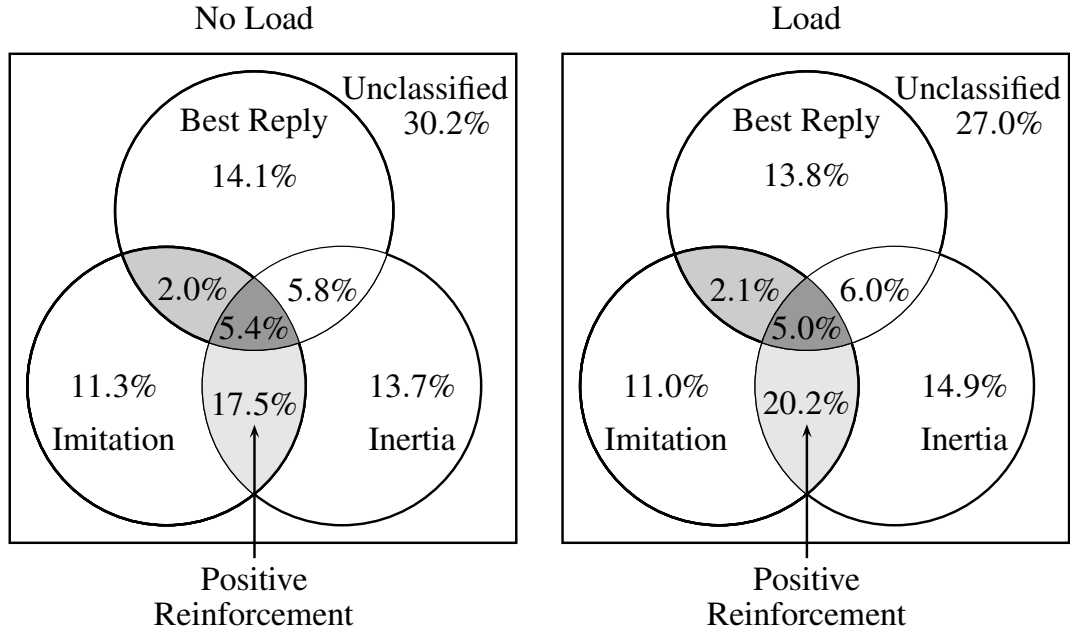


Figure 3: Overview of Observations and their Classification, Replications.
Note: Classification of decisions in NoLoad (left-hand side) and Load (right-hand side).

Of those, 23.65 percent were myopic best replies and 37.28 percent were imitative. The remaining (16.32 percent) were under alignment, of which 43.79 percent were myopic best replies (hence also imitative).

4.2 Results

We now proceed to examine all our previous hypotheses, (H1) to (H4b), separately for both replications. Figure 4 depicts the averages of individual average response times conditional on imitative choices, myopic best replies, and other choices for the NoLoad and Load replications, hence illustrating the first three hypotheses. (H1) states that decisions favored by the most deliberative rule, in our case myopic best reply, must be slower in case of conflict with imitation than in case of alignment. Indeed, myopic best replies are significantly slower under conflict than under alignment both in NoLoad ($p = .0277$) and in Load ($p = .0464$). (H2) states that, in case of conflict, myopic best replies must be slower than imitative decisions. This is the case both for NoLoad and for Load ($p < .03$).¹¹ (H3) predicts that, in case of alignment, myopic best replies (which are also imitative in this case) should be faster than other decisions. Again, we find significant differences in the predicted direction both in NoLoad and in Load ($p < .03$).

Figure 5 shows the relative frequencies of different types of decisions in conflict and alignment for the NoLoad (left-hand side) and Load (right-hand side) replications, illustrating (H4a) and (H4b). As predicted by (H4a), myopic best replies are significantly

¹¹The replications had less participants and a larger block size which resulted in fewer number of observations for this analysis. Most paired block-level tests yield the smallest possible p -value which is .0277 for $N = 6$.

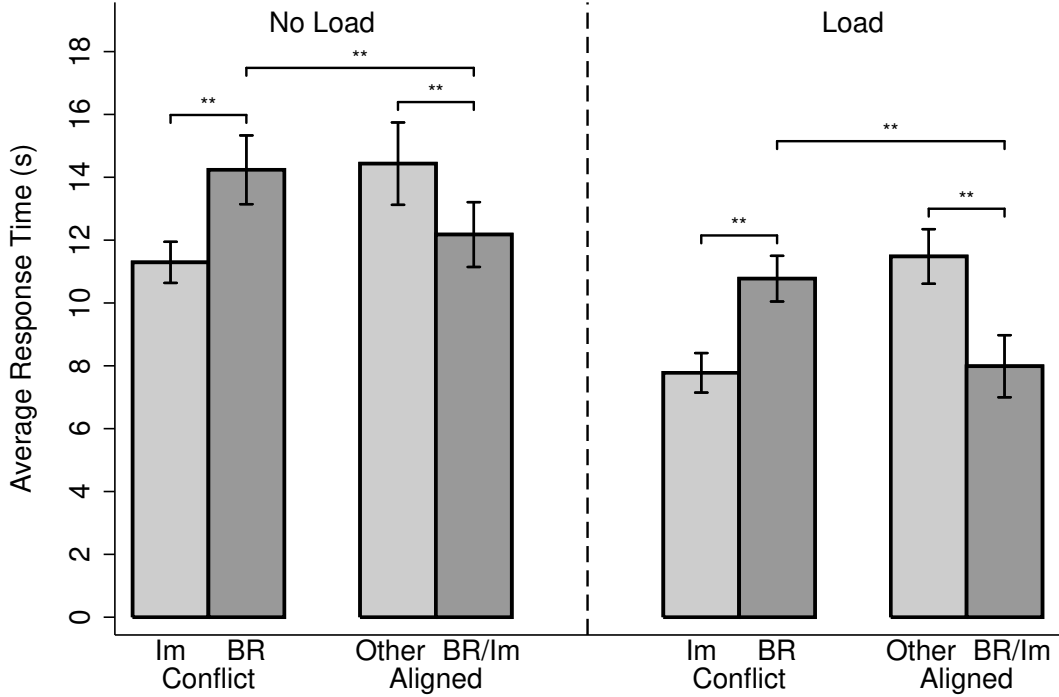


Figure 4: Average Response Times, Replications.

Note: Average response times of imitative decisions, myopic best replies, and other choices in the NoLoad (left-hand side) and Load (right-hand side) replications. Stars indicate Wilcoxon Signed-Rank tests. ** $p < .05$. Note that $p < .01$ is not possible for WSR tests with 6 block-level paired observations.

less frequent in conflict than in alignment, both in NoLoad and in Load ($p < .03$). Hypothesis (H4b) is also clearly confirmed by the data in both treatments (we remark that this was the only hypothesis not strongly supported in the main experiment). That is, imitative decisions were significantly less frequent in conflict than in alignment both in NoLoad (conflict, average 34.65 percent; alignment, 43.55; WSR, $N = 6$, $z = -2.201$, $p = .0277$) and in Load replications (conflict, 37.17 percent; alignment, 44.02; $N = 6$, $z = -1.992$, $p = .0464$).¹²

In summary, conservative, two-tailed nonparametric tests at the block level confirm all of the model's predictions for both replications, showing that the results obtained in main experiment were robust. Analogously to Section 3.3, we now conduct panel regressions making use of all $48 \times 72 = 3,456$ observations of each replication.

Table 5 presents random effects panel regressions of the log-transformed response times. For conciseness, we present only the analogues of Models 1 and 3 from the main experiment for each replication (recall Table 2).¹³

¹²In the Load replication, one group successfully colluded in the last part of the experiment. Excluding the corresponding block yields comparable results.

¹³Model 2 does not qualitatively differ from the other models.

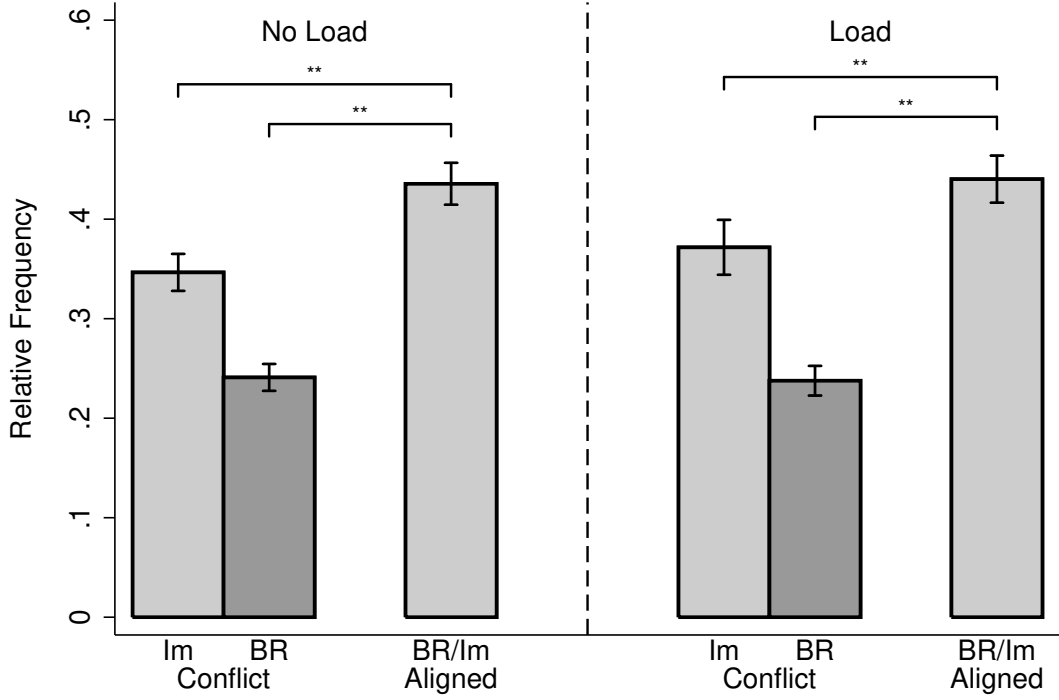


Figure 5: Relative Frequency of Imitation and Myopic Best Reply, Replications.
Note: Relative frequencies of imitative decisions (Im) and myopic best replies (BR) in the NoLoad (left-hand side) and Load replications (right-hand side).

Table 5: Random Effects Panel Regressions on (log) Response Times, Replications.

	No Load		Load	
ln(ResponseTime)	Model 1	Model 3	Model 1	Model 3
Conflict	0.1505*** (0.0233)	0.1470*** (0.0175)	0.1900*** (0.0605)	0.1937*** (0.0540)
Imitation-Conflict	-0.2337*** (0.0235)	-0.2172*** (0.0181)	-0.2759*** (0.0556)	-0.2258*** (0.0321)
Other	0.1619*** (0.0345)	0.1255*** (0.0347)	0.2559*** (0.0607)	0.2261*** (0.0468)
Other × Conflict	-0.1688*** (0.0479)	-0.1310*** (0.0457)	-0.2772*** (0.0658)	-0.2497*** (0.0550)
Collusion				-0.0739 (0.0825)
Constant	2.3230*** (0.0637)	3.0231*** (0.1948)	1.9410*** (0.0838)	1.7145*** (0.1749)
Controls	No	Yes	No	Yes
Demographics	No	Yes	No	Yes
R^2	0.0348	0.1216	0.0623	0.1559

Note: Standard errors, clustered by 12 matching blocks, in parentheses.

Model 1 in Table 5 tests for the basic effects. Model 3 adds a Collusion dummy (only relevant for Load) and further Controls and Demographics, to show the robustness

Table 6: Probit Regression Models for Myopic Best Reply, Replications.

	No Load		Load	
Myopic Best Reply	Model 1	Model 3	Model 1	Model 3
Conflict	−0.5615*** (0.0832)	−0.5737*** (0.0717)	−0.5491*** (0.0648)	−0.5425*** (0.0653)
Collusion				−0.1059*** (0.0379)
Constant	−0.1583*** (0.0488)	−0.6876*** (0.2356)	−0.1796*** (0.0517)	0.0299 (0.2267)
Controls	No	Yes	No	Yes
Demographics	No	Yes	No	Yes
Log Pseudolikelihood	−1976.1410	−1964.1982	−1956.3004	−1952.2798

Note: Standard errors, clustered by 12 matching blocks, in parentheses.

of the results. The coefficient for Conflict is positive and highly significant (NoLoad, $p < .0001$ in both models; Load, Model 1 $p = .0017$, Model 3 $p = .0003$), showing that myopic best replies are slower under conflict than under alignment, as predicted by (H1). As in Table 2, the coefficient for Imitation-Conflict captures the difference between imitative decisions and myopic best replies in case of conflict. The coefficient is negative and highly significant ($p < .0001$ in all models), confirming (H2). The coefficient for Other decisions captures the difference, in case of alignment, between other decisions and myopic best replies (which are also imitative). The coefficient is positive and highly significant (NoLoad, Model 1 $p < .0001$, Model 3 $p = .0003$; Load, $p < .0001$ in both models), confirming (H3).

As in the main experiment, we also examine a few exploratory questions. The linear combination of the coefficients for Other and Other×Conflict is not significantly different from zero, indicating that (as in the main experiment) in case of conflict myopic best replies are not faster than other decisions. In contrast, and again as in the main experiment, a linear combination test shows that imitation decisions are significantly faster than other decisions in conflict ($p < .0001$ in all models), which might indicate that part of the latter decisions reflect more complex behavioral rules. The Collusion dummy, identifying subjects who colluded, was not significant (in contrast to the main experiment).

Tables 6 and 7 present probit panel regressions with myopic best reply and imitation as dependent variable, respectively. Both reveal a highly-significant and negative Conflict dummy (Best Reply, Table 6, all $p < .0001$; Imitation, Table 7, No Load: Model 1 $p = .0019$, Model 3 $p = .0009$, Load: Model 1 $p = .0157$, Model 3 $p = .0008$). This confirms that, in both replications, both best myopic best replies and imitative choices were significantly less frequent under conflict than under alignment, as predicted in (H4a) and (H4b), respectively.

In summary, the regression analysis confirms all our conclusions from the main experiment in both replications and hence supports the interpretation that decisions arise

Table 7: Probit Regression Models for Imitation, Replications.

	No Load		Load	
Imitation	Model 1	Model 3	Model 1	Model 3
Conflict	−0.2190*** (0.0707)	−0.2330*** (0.0700)	−0.1423** (0.0589)	−0.1800*** (0.0535)
Collusion				0.4186 (0.3999)
Constant	−0.1935*** (0.0557)	−0.0118 (0.3526)	−0.2076*** (0.0656)	0.1838 (0.5511)
Controls	No	Yes	No	Yes
Demographics	No	Yes	No	Yes
Log Pseudolikelihood	−2121.5558	−2111.0881	−2054.3477	−2016.3972

Note: Standard errors, clustered by 12 matching blocks, in parentheses.

from the interaction of two clearly-differentiated behavioral rules along the lines of myopic best reply and imitation of best outcomes.

5 Other Behavioral Rules

Two further behavioral rules can be defined and analyzed with our data. The first is a simple “win-stay” version of reinforcement learning, i.e. the tendency to repeat what has worked in the past without paying attention to whether the conditions in which past actions were successful have changed. Reinforcement is particularly important for economics, as it captures the empirically-relevant focus on *past performance*, whose consequences are well-documented (e.g., outcome bias; Baron and Hershey, 1988; Dillon and Tinsley, 2008). Evidence from neuroscience has shown that reinforcement learning is associated with extremely fast and unconscious brain responses (e.g., Schultz, 1998; Holroyd and Coles, 2002). In an explicitly economic context, Achtziger and Alós-Ferrer (2014) showed that a simple reinforcement heuristic corresponds to a highly automatic process which competes with more deliberative rules when feedback comes in a win-loss frame.

In a Cournot oligopoly, “win-stay” corresponds to *positive reinforcement*, which prescribes to repeat the previous choice if the player has “won,” that is, obtained the maximum observed profits. This rule can be described as “imitating yourself” and coincides with imitation in a subset of decision situations (recall Figures 1 and 3). The remaining imitative decisions correspond to imitating others. Since reinforcement is considered to be based on rather automatic processes, we hypothesize that it should lead to shorter response times than “imitating others.”¹⁴

¹⁴Note that imitating yourself and imitating others are never simultaneously active processes, but rather constitute a partition of imitative decisions and hence the prediction of faster response times is straightforward: the process favoring imitation is faster in one case than in the other, while the competing myopic best reply rule remains fixed.

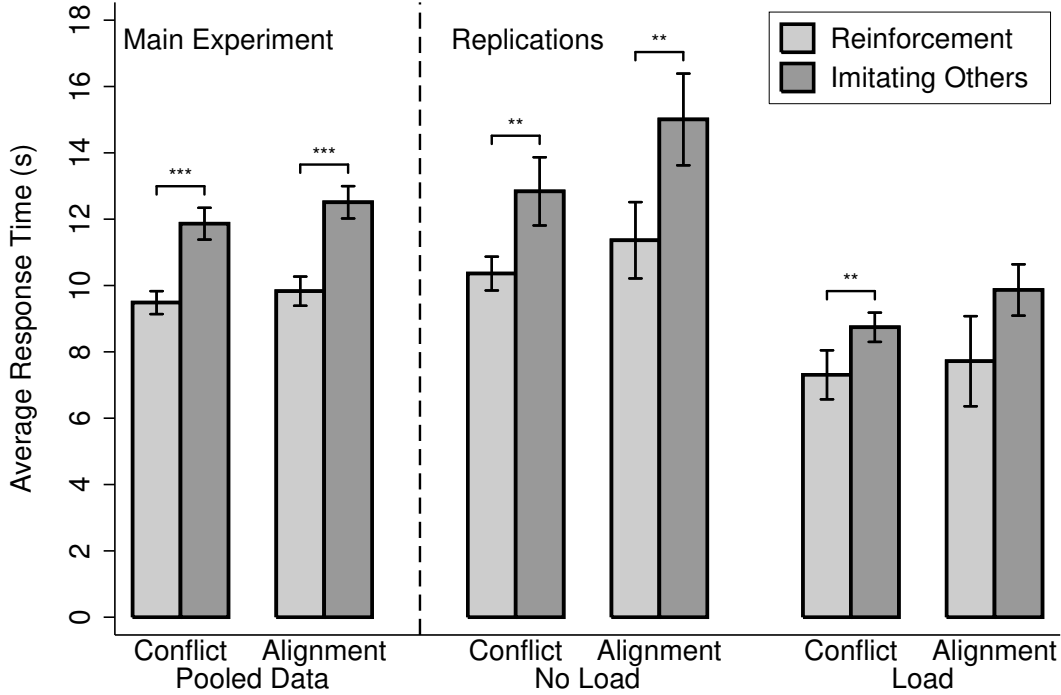


Figure 6: Average Response Times of Reinforcement and Imitating-others Decisions
Note: Left-hand side: Main Experiment. Right-hand side: Replications. Stars indicate Wilcoxon Signed-Rank tests. ** $p < .05$ and *** $p < .01$.

Figure 6 displays the response times of decisions where participants imitated themselves or others, both for the main experiment (left-hand side) and both replications (right-hand side). For completeness, we disentangle the comparison according to whether imitation (or positive reinforcement) was in conflict or in alignment with myopic best reply. Reinforcement decisions in case of conflict were significantly faster than imitating-others decisions in all cases (main experiment: average 9.48 s vs. 11.86 s, WSR, $N = 16$, $z = -3.258$, $p = .0011$; Replications, NoLoad; 10.36 s vs. 12.84 s, $N = 6$, $z = -2.201$, $p = .0277$; Replications, Load: 7.31 s vs. 8.74 s, $N = 6$, $z = -1.992$, $p = .0464$). In case of alignment, reinforcement decisions were also significantly faster than imitating-others decisions in the main experiment (average 9.62 s vs. 13.71 s; $N = 16$, $z = -2.947$, $p = .0032$) and NoLoad replication (11.36 s vs. 15.01 s; $N = 6$, $z = -1.992$, $p = .0464$). In the Load replication reinforcement decisions were also faster than imitating-others decisions, but there were no significant differences (7.72 s vs. 9.87 s, $N = 6$, $z = -1.153$, $p = .2489$).

In summary, we confirm that the imitation behavioral rule that we consider might be supported by a composite process which, in some cases, reflects positive reinforcement. This is of independent interest, but does not change our previous conclusions.

Decisions following positive reinforcement (or imitating yourself) imply upholding the previously-selected action. Hence, they are aligned with a further, particularly simple

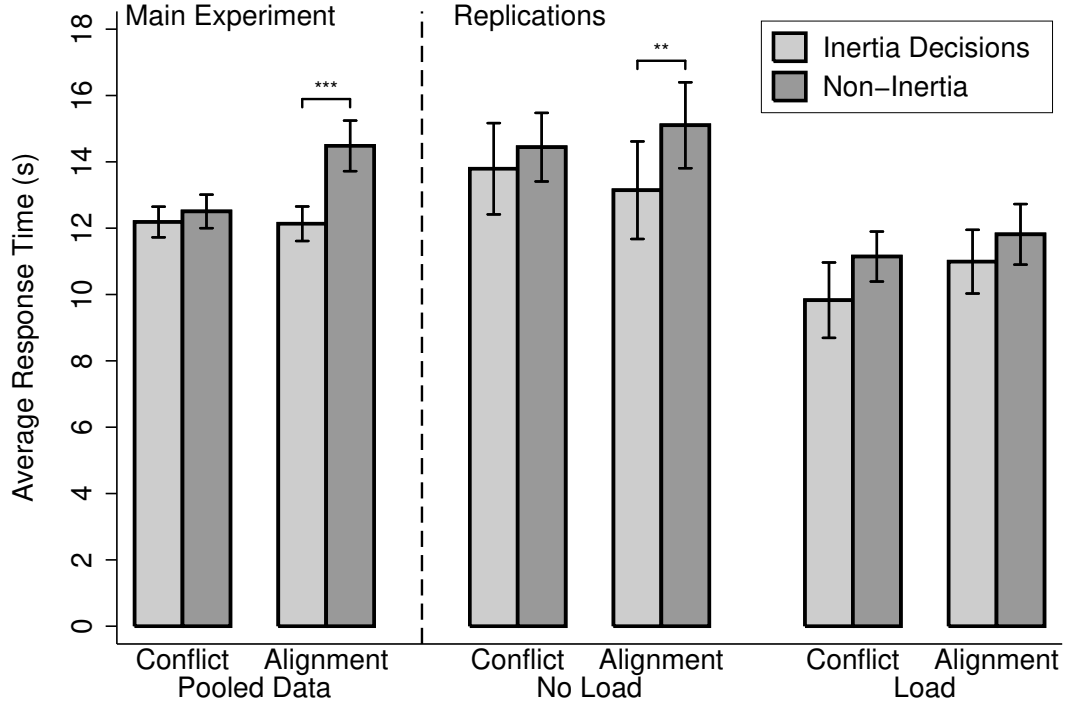


Figure 7: Average Response Times of Stay and Shift Decisions

Note: Left-hand side: Main Experiment. Right-hand side: Replications. Each subset shows stay and shift best reply decisions in case of conflict and other (non-best replies) stay and shift decisions in case of alignment. Stars indicate Wilcoxon Signed-Rank tests. ** $p < .05$ and *** $p < .01$.

behavioral rule: *decision inertia*, i.e. the tendency to repeat previous behavior independently of any feedback. This raises the natural question of whether the driver of the effects above is actually this simple but more general rule, i.e. whether inertia results in clear effects beyond situations where the decision maker has obtained the largest profits. Previous work (Alós-Ferrer et al., 2016) has compared decision inertia with reinforcement in the belief-updating task of Charness and Levin (2005) and Achtziger and Alós-Ferrer (2014), and found that inertia does cause asymmetries in error rates, but this behavioral rule seems weaker than reinforcement and is typically washed away by it. To see whether inertia is behaviorally relevant in our paradigm, we examined it in the cases where it is *not* aligned with imitation, since in case of alignment of imitation we obtain positive reinforcement. To avoid confusion, however, we reserve the words “alignment” and “conflict” for the confluence or not of myopic best reply and imitation. In case of conflict between myopic best reply and imitation, we will compare “stay” myopic best replies (as prescribed with inertia) with “shift” myopic best replies. In case of alignment between myopic best reply and imitation, we test within other kind of decisions not following the common prescription of myopic best reply and imitation.

Figure 7 depicts the response times of decisions in line with inertia (“stay” decisions) and those opposed to it (“shift” decisions), both for the main experiment (left-hand side) and for both replications (right-hand side). For conflict, the comparison is between “stay” and “shift” best replies. There were, however, no differences in the response times of these two kinds of decisions, neither in the main experiment (stay, average 12.19 s; shift, 12.51 s; WSR, $N = 16$, $z = -0.776$, $p = .4380$) nor in any replication (NoLoad, 13.79 s vs. 14.44 s, $N = 6$, $z = -1.153$, $p = .2489$; Load, 9.83 s vs. 11.14 s, $N = 6$, $z = -0.943$, $p = .3454$). Hence, whenever myopic best reply and imitation conflict, there is no evidence of involvement of inertia (beyond the possible confluence with imitation), and in particular the effects of reinforcement described in the previous subsection are unlikely to be due to a more general process reflecting pure inertia.

For alignment (between imitation and myopic best reply), we compare all *non*-best replies of the “stay” and “shift” forms. Such stay (inertia) decisions were significantly faster than the comparable shift decisions in the main experiment (stay, average 12.13 s; shift, 14.48 s; WSR, $N = 16$, $z = -2.741$, $p = .0061$) and in NoLoad replication (13.14 s vs. 15.10 s, $N = 6$, $z = -1.992$, $p = .0464$). There were no significant differences in the Load replication (10.99 s vs. 11.81 s, $N = 6$, $z = -1.153$, $p = .2489$). These results are interesting. In this case, best replies coincide with imitative decisions, that is, the “other” decisions we examine are not imitative, and in particular can *not* follow from positive reinforcement. Although this is speculative, these results suggest that shift decisions in this case might include choices derived from higher-order reasoning or more complex behavioral rules. This would be consistent with the long response times of “other” decisions under conflict discussed in the regression analyses in Sections 3.3 and 4.2.

6 Conclusion

Economics has long embraced the idea that human decision makers have limited capacity and hence will display bounded rationality (Simon, 1959). There is little doubt that humans rely on cognitive shortcuts, often unconsciously (Kahneman, 2003). Microeconomic theory has captured such ideas through behavioral rules reflecting different degrees of sophistication (e.g. Samuelson, 1997; Fudenberg and Levine, 1998), which can often be usefully classified along the lines sketched in dual-process models from psychology (Weber and Johnson, 2009). Other systematic attempts, as prospect theory (Kahneman and Tversky, 1979) or quasi-hyperbolic discounting (Laibson, 1997), have incorporated these ideas in the form of specific modifications of standard economic models as expected utility or exponential discounting. Although such models have clearly shown its usefulness, they remain *as if* in the sense that they implicitly or explicitly assume a multiplicity of behavioral rules at the individual level, but deliver no explicit means to test for that multiplicity.

We provide a simple formal model where economic agents decide following different behavioral rules which differ along the cognitive dimension. The model makes a number of non-trivial predictions which allow us to directly test for the multiplicity of behavioral rules. This is possible because the predictions rely on explicit characteristics of the behavioral rules which reach beyond the *as if* description needed to fit choices. First, predictions rely on the *ex ante* classification of decisions in *conflict* or *alignment* according to the pre-specified behavioral rules. Second, they concern both choices *and* response times, the latter being a direct correlate of the postulated characteristics of the brain decision processes underlying the behavioral rules.

In one main experiment and two replications, we find overwhelming evidence in favor of the presence of multiple behavioral rules. We focus on an economically-relevant setting, a Cournot oligopoly where myopic best reply is the natural candidate for a more deliberative behavioral rule, but previous research has identified imitation of best outcomes as a relevant, more impulsive competing rule, both theoretically (Vega-Redondo, 1997; Alós-Ferrer and Ania, 2005) and experimentally (Huck et al., 1999, 2004; Offerman et al., 2002; Apesteguía et al., 2007). We find a number of “smoking guns,” all predicted by our model, which point to the multiplicity of behavioral rules: best replies are slower under conflict with imitation than under alignment (generalizing the Stroop effect from cognitive psychology), they are slower than imitative decisions under conflict but slower than other decisions under alignment, and both best replies and imitative decisions are less frequent under conflict than under alignment. The evidence is striking and systematic, and speaks in favor of a *literal* multiplicity of competing behavioral rules in economic decision making.

In conclusion, our model and empirical evidence strongly suggest that economic decision making can often be better explained by integrating different views of behavior, instead of either assuming fully-rational optimization or boundedly-rational impulse-response behavior only. Multiple behavioral rules are more than a convenient metaphor or an *as if* model, and economic modeling can be greatly improved by viewing decisions as the result of the interaction of different behavioral rules and decision processes in the human brain.

References

- Achtziger, A. and C. Alós-Ferrer (2014). Fast or Rational? A Response-Times Study of Bayesian Updating. *Management Science* 60(4), 923–938.
- Achtziger, A., C. Alós-Ferrer, S. Hügelschäfer, and M. Steinhäuser (2014). The Neural Basis of Belief Updating and Rational Decision Making. *Social Cognitive and Affective Neuroscience* 9(1), 55–62.
- Achtziger, A., C. Alós-Ferrer, and A. Ritschel (2019). Cognitive Load in Economic Decisions: How To Tell Whether It Works. Working Paper, University of Zurich.

- Alós-Ferrer, C. (2018). A Dual-Process Diffusion Model. *Journal of Behavioral Decision Making* 31(2), 203–218.
- Alós-Ferrer, C. and A. B. Ania (2005). The Evolutionary Stability of Perfectly Competitive Behavior. *Economic Theory* 26, 179–197.
- Alós-Ferrer, C. and J. Buckenmaier (2018). Cognitive Sophistication and Deliberation Times. Working Paper, University of Zurich.
- Alós-Ferrer, C., D.-G. Granić, J. Kern, and A. K. Wagner (2016). Preference Reversals: Time and Again. *Journal of Risk and Uncertainty* 52(1), 65–97.
- Alós-Ferrer, C., S. Hügelschäfer, and J. Li (2016). Inertia and Decision Making. *Frontiers in Psychology* 7 (169), 1–9.
- Alós-Ferrer, C. and A. Ritschel (2018). The Reinforcement Heuristic in Normal Form Games. *Journal of Economic Behavior and Organization* 152, 224–234.
- Alós-Ferrer, C. and F. Strack (2014). From Dual Processes to Multiple Selves: Implications for Economic Behavior. *Journal of Economic Psychology* 41, 1–11.
- Apesteguía, J., S. Huck, and J. Oechssler (2007). Imitation—Theory and Experimental Evidence. *Journal of Economic Theory* 136, 217–235.
- Apesteguía, J., S. Huck, J. Oechssler, and S. Weidenholzer (2010). Imitation and The Evolution of Walrasian Behavior: Theoretically Fragile But Behaviorally Robust. *Journal of Economic Theory* 145, 1603–1617.
- Baddeley, A. D. (1992). Working Memory. *Science* 255(5044), 556–559.
- Baddeley, A. D., D. Chincotta, and A. Adlam (2001). Working Memory and the Control of Action: Evidence From Task Switching. *Journal of Experimental Psychology: General* 130(4), 641–657.
- Bargh, J. A. (1989). Conditional Automaticity: Varieties of Automatic Influences in Social Perception and Cognition. In J. S. Uleman and J. A. Bargh (Eds.), *Unintended Thought*, pp. 3–51. Guilford.
- Baron, J. and J. C. Hershey (1988). Outcome Bias in Decision Evaluation. *Journal of Personality and Social Psychology* 54(4), 569–579.
- Blume, L. (1993). The Statistical Mechanics of Strategic Interaction. *Games and Economic Behavior* 5, 387–424.
- Carpenter, J., M. Graham, and J. Wolf (2013). Cognitive ability and strategic sophistication. *Games and Economic Behavior* 80, 115–130.
- Charness, G. and D. Levin (2005). When Optimal Choices Feel Wrong: A Laboratory Study of Bayesian Updating, Complexity, and Affect. *American Economic Review* 95(4), 1300–1309.
- De Neys, W., O. Vartanian, and V. Goel (2008). Smarter than We Think: When Our Brains Detect That We Are Biased. *Psychological Science* 19(5), 483–489.

- Dillon, R. L. and C. H. Tinsley (2008). How Near-Misses Influence Decision Making Under Risk: A Missed Opportunity for Learning. *Management Science* 54(8), 1425–1440.
- Erev, I. and A. E. Roth (1998). Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria. *American Economic Review* 88(5), 848–881.
- Evans, J. S. (2008). Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition. *Annual Review of Psychology* 59, 255–278.
- Fischbacher, U. (2007). z-Tree: Zurich Toolbox for Ready-Made Economic Experiments. *Experimental Economics* 10(2), 171–178.
- Fischbacher, U., R. Hertwig, and A. Bruhin (2013). How to Model Heterogeneity in Costly Punishment: Insights from Responders’ Response Times. *Journal of Behavioral Decision Making* 26(5), 462–476.
- Fudenberg, D. and D. Levine (1998). *The Theory of Learning in Games*. Cambridge, Massachusetts: The MIT Press.
- Fudenberg, D., P. Strack, and T. Strzalecki (2018). Speed, Accuracy, and the Optimal Timing of Choices. *American Economic Review* 108(12), 3651–3684.
- Greiner, B. (2015). Subject Pool Recruitment Procedures: Organizing Experiments with ORSEE. *Journal of the Economic Science Association* 1, 114–125.
- Holroyd, C. B. and M. G. Coles (2002). The Neural Basis of Human Error Processing: Reinforcement Learning, Dopamine, and the Error-Related Negativity. *Psychological Review* 109, 679–709.
- Huck, S., H.-T. Normann, and J. Oechssler (1999). Learning in Cournot Oligopoly - An Experiment. *Economic Journal* 109, C80–C95.
- Huck, S., H.-T. Normann, and J. Oechssler (2004). Through Trial and Error to Collusion. *International Economic Review* 45(1), 205–224.
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *American Economic Review* 93(5), 1449–1475.
- Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision Under Risk. *Econometrica* 47(2), 263–291.
- Kandori, M., G. J. Mailath, and R. Rob (1993). Learning, Mutation, and Long Run Equilibria in Games. *Econometrica* 61, 29–56.
- Krajbich, I., B. Bartling, T. Hare, and E. Fehr (2015). Rethinking Fast and Slow Based on a Critique of Reaction-Time Reverse Inference. *Nature Communications* 6(7455), 1–9.
- Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics* 2(112), 443–477.
- MacCleod, C. M. (1991). Half a Century of Research on the Stroop Effect: An Integrative Review. *Psychological Bulletin* 109(2), 163–203.

- Moffatt, P. G. (2005). Stochastic Choice and the Allocation of Cognitive Effort. *Experimental Economics* 8(4), 369–388.
- Nieuwenhuis, S., N. Yeung, W. van den Wildenberg, and K. R. Ridderinkhof (2003). Electrophysiological Correlates of Anterior Cingulate Function in a Go/No-Go Task: Effects of Response Conflict and Trial Type Frequency. *Cognitive, Affective, and Behavioral Neuroscience* 3(1), 17–26.
- Offerman, T., J. Potters, and J. Sonnemans (2002). Imitation and Belief Learning in an Oligopoly Experiment. *Review of Economic Studies* 69(4), 973–997.
- Ratcliff, R. (1978). A Theory of Memory Retrieval. *Psychological Review* 85, 59–108.
- Ratcliff, R. and J. N. Rouder (1998). Modeling Response Times for Two-Choice Decisions. *Psychological Science* 9(5), 347–356.
- Rubinstein, A. (2007). Instinctive and Cognitive Reasoning: A Study of Response Times. *Economic Journal* 117(523), 1243–1259.
- Rubinstein, A. (2016). A Typology of Players: Between Instinctive and Contemplative. *Quarterly Journal of Economics* 131(2), 859–890.
- Samuelson, L. (1997). *Evolutionary Games and Equilibrium Selection*. Cambridge, Massachusetts: The MIT Press.
- Schaffer, M. E. (1989). Are Profit-Maximisers the Best Survivors? *Journal of Economic Behavior and Organization* 12, 29–45.
- Schultz, W. (1998). Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology* 80(1), 1–27.
- Shadlen, M. N. and D. Shohamy (2016). Decision Making and Sequential Sampling from Memory. *Neuron* 90, 927–939.
- Simon, H. A. (1959). Theories of Decision Making in Economics. *American Economic Review* 49(1), 223–283.
- Spiliopoulos, L. (2018). The Determinants of Response Time in a Repeated Constant-sum Game: A Robust Bayesian Hierarchical Dual-Process Model. *Cognition* 172, 107–123.
- Spiliopoulos, L. and A. Ortmann (2018). The BCD of Response Time Analysis in Experimental Economics. *Experimental Economics* 21(2), 383–433.
- Stroop, J. R. (1935). Studies of Interference in Serial Verbal Reactions. *Journal of Experimental Psychology* 35, 643–662.
- Vega-Redondo, F. (1997). The Evolution of Walrasian Behavior. *Econometrica* 65(2), 375–384.
- Weber, E. U. and E. J. Johnson (2009). Mindful Judgment and Decision Making. *Annual Review of Psychology* 60, 53–85.

Appendix

A.1 Proofs

Proof of Theorem 1. The expected response time of best replies in case of alignment ($x^B = x^I$) is

$$E(RT|x^B, \text{Alignment}) = \frac{(1 - \Delta)P^{BR}R^B + \Delta P^{Im}R^I}{(1 - \Delta)P^{BR} + \Delta P^{Im}}$$

and the expected response time of best replies in case of conflict ($x^B \neq x^I$) is

$$E(RT|x^B, \text{Conflict}) = \frac{(1 - \Delta)P^{BR}R^B + \Delta P_B^{Im}R^I}{(1 - \Delta)P^{BR} + \Delta P_B^{Im}}.$$

Then, $E(RT|x^B, \text{Conflict}) > E(RT|x^B, \text{Alignment})$ holds if and only if

$$(P^{Im} - P_B^{Im})R^B > (P^{Im} - P_B^{Im})R^I$$

which holds by (P1) and (R). □

Proof of Theorem 2. (H2) The expected response time of best replies in case of conflict ($x^B \neq x^I$) is as given in the proof of Theorem 1, and the expected response time of imitative answers is

$$E(RT|x^I, \text{Conflict}) = \frac{(1 - \Delta)P_I^{BR}R^B + \Delta P^{Im}R^I}{(1 - \Delta)P_I^{BR} + \Delta P^{Im}}.$$

where P_I^{BR} denotes the probability with which the best reply rule selects an imitative answer when it does not coincide with the prescription of imitation, i.e. $P_I^{BR} = P^{BR}(x^I)$ when $x^B \neq x^I$. Then, $E(RT|x^B, \text{Conflict}) > E(RT|x^I, \text{Conflict})$ if and only if

$$(P^{BR}P^{Im} - P_I^{BR}P_B^{Im})(R^B - R^I) > 0$$

which holds by (R) ($R^B > R^I$) and (P1) (which implies $P^{BR}P^{Im} > P_I^{BR}P_B^{Im}$).

(H3) The expected response time of best replies in case of alignment ($x^B = x^I$) is as given in the proof of Theorem 1, and the expected response time of other answers is

$$E(RT|x \neq x^B, \text{Alignment}) = \frac{(1 - \Delta)(1 - P^{BR})R^B + \Delta(1 - P^{Im})R^I}{(1 - \Delta)(1 - P^{BR}) + \Delta(1 - P^{Im})}.$$

Then, $E(RT|x^B, \text{Alignment}) < E(RT|x \neq x^B, \text{Alignment})$ if and only if

$$((1 - P^{BR})P^{Im} - P^{BR}(1 - P^{Im}))(R^B - R^I) > 0.$$

Since $R^B > R^I$ holds by (R), the result holds if

$$(1 - P^{BR})P^{Im} > P^{BR}(1 - P^{Im})$$

which is equivalent to $P^{Im} > P^{BR}$. The latter holds by (P2). □

Proof of Theorem 3. (H4a) The proportion of best replies in case of alignment ($x^B = x^I$) is

$$P(BR|Alignment) = (1 - \Delta)P^{BR} + \Delta P^{Im}$$

and the proportion of best replies in case of conflict ($x^B \neq x^I$) is

$$P(BR|Conflict) = (1 - \Delta)P^{BR} + \Delta P_B^{Im}$$

where P_B^{Im} denotes the probability with which the imitation rule selects a best reply when it does not coincide with the prescription of imitation, i.e. $P_B^{Im} = P^{Im}(x^B)$ when $x^B \neq x^I$. Then, $P(BR|Alignment) > P(BR|Conflict)$ if and only if $P^{Im} > P_B^{Im}$, which holds by (P1).

(H4b) is analogous to (H4a). \square

Proof of Theorem 4. A straightforward computation analogous to the proof of Theorem 1 shows that (H1) reduces to

$$(t_C - t_A) + [(1 - \Delta_C)\Delta_A P^{Im} - (1 - \Delta_A)\Delta_C P_B^{Im}] (R^B - R^I) > 0.$$

Since $t_C \geq t_A$ and $R^B > R^I$ by (R), this follows if the left-hand bracket is positive, or, equivalently,

$$\frac{\Delta_A}{1 - \Delta_A} P^{Im} > \frac{\Delta_C}{1 - \Delta_C} P_B^{Im}$$

which follows from (P1) and $\Delta_A \geq \Delta_C$.

(H2) and (H3) are unaffected by the extension, since they refer either to the case of conflict or the case of alignment, and not to comparisons across them. The proof of Theorem 2 goes through replacing Δ with the corresponding Δ_i .

(H4a) and (H4b) do not refer to response times, hence they are unaffected by non-decision times. The assumption that $\Delta_A \geq \Delta_C$, however, does influence the proportion of choices of each type. A straightforward computation analogous to the proof of Theorem 3 shows that (H4a) is equivalent to

$$\Delta_A P^{Im} - \Delta_C P_B^{Im} > (\Delta_A - \Delta_C) P^{BR}.$$

By (P1), $\Delta_A P^{Im} - \Delta_C P_B^{Im} > (\Delta_A - \Delta_C) P^{Im}$. The latter is larger than $(\Delta_A - \Delta_C) P^{BR}$ by (P2) and $\Delta_A \geq \Delta_C$.

A similar computation shows that (H4b) is equivalent to

$$(\Delta_A - \Delta_C) P^{Im} + [(1 - \Delta_A) P^{BR} - (1 - \Delta_C) P_I^{BR}] > 0$$

Since $P^{Im} > P^{BR}$ by (P2) and $\Delta_A \geq \Delta_C$, the left-hand side is larger than $(1 - \Delta_C) (P^{BR} - P_I^{BR})$, which is strictly larger than zero by (P1). This completes the proof. \square

A.2 Translated Instructions

General Instructions

The experiment consists of **three parts with 17 rounds each** in which **you and three other participants** make decisions. After the completion of these three parts, a questionnaire will follow. In each of the three parts you will **earn points**. How many points you earn depends on your decisions and the decisions of the players in your group. All points you earned each round will be added up at the end of the experiment and exchanged into Euros. The exchange rate is:

Main Experiment: 1000 points= 18 Eurocents.

Replications: 1000 points= 20 Eurocents.

Independently of your decisions, you will receive 2.50 EUR for your participation. The total amount will be paid **in cash** and **anonymously** at the end of the experiment.

On the following pages you will receive all further information which you need for the experiment. Among other things the sequence of the experiment will be explained in detail. Once you have finished reading the instructions, please proceed to answer the control questions on the screen.

Instructions of the Experiment

General Sequence: The experiment is divided into **three parts**. The **procedure is the same** for each part. Only the payoff table (which will be discussed later in more detail) and the composition of the groups change with each part. One part consists of **17 rounds**. At the beginning of each part, participants will be divided into groups. One **group** consists of **4 players** (you included) and stays the same for the duration of a part. That means, that you always interact with the same players during one part. In every new part two of players will be replaced and therefore the composition of the group **changes**. That means that in a new part you do not interact with the same players as in the previous part. In every round you have to decide among four options, *A*, *B*, *C*, or *D*. How many points you earn in one round depends both on your choice and on the choices of the other three group members. *[Load Replication:]* In addition you can earn additional points in every round.

Payoff Tables: The payoff tables are an important component of the experiment. They show you all possible payoffs depending on your choice and the choice of the other three group members. The rows represent your choice and the columns represent the joint choice of the other group members. The appropriate cell entry is the amount of points you would receive if this combination of choices occurs. Please note that for your payoff it is irrelevant which of the other group members made which choice. That means that if the other group members choose *C*, *A*, and *B*, respectively, this has the same effect on your payoff than if they choose *A*, *B*, and *C*. For a better overview, columns are ordered alphabetically.

Figures A.1-A.3/*[Load Replication:]*A.1-A.5 depict examples of such payoff tables. Please note that in the experiment other payoff tables will be used. **Important note:** The payoff table will not change during a part. The same payoff table applies to all group members.

Your Decision: In each round you have to choose one of the four options, *A*, *B*, *C*, or *D*. You have 30 seconds to make your choice. You make a choice by clicking on the appropriate button on the screen. During your choice the payoff table of the current

part will be shown. The next round begins as soon as all participants made their choice. In every round the result of the previous round is shown (except for the first round of every part).

Sequence of Decisions in a Round in Detail: The **payoff table** will be shown at the beginning of each part so you can familiarize yourself with it (see Figure A.1). The table will be kept on the screen during the experiment at all times – you **do not** have to memorize or copy the table. After you have familiarized yourself with the table click “continue”. The **decision phase** will start as soon as all participants are ready.

Round 1 of 51																				
	AAA	AAB	AAC	AAD	ABB	ABC	ABD	ACC	ACD	ADD	BBB	BBC	BBD	BCC	BCD	BDD	CCC	CCD	CDD	DDD
A	1000	1200	750	825	1400	950	1025	500	575	650	1600	1150	1225	700	775	850	250	325	400	475
B	816	952	646	697	1088	782	833	476	527	578	1224	918	969	612	663	714	306	357	408	459
C	1050	1330	700	805	1610	980	1085	350	455	560	1890	1260	1365	630	735	840	0	105	210	315
D	1056	1312	736	832	1568	992	1088	416	512	608	1824	1248	1344	672	768	864	96	192	288	384

For the next 17 rounds (1-17) this will be the payoff table.

Click "continue" so the experiment can continue.

continue

Figure A.1: Beginning of a part.

Now you can choose among **four options**, *A*, *B*, *C*, and *D*. To make a choice click on the appropriate button of your choice (see Figure A.2).

Round 1 of 51																				
	AAA	AAB	AAC	AAD	ABB	ABC	ABD	ACC	ACD	ADD	BBB	BBC	BBD	BCC	BCD	BDD	CCC	CCD	CDD	DDD
A	1000	1200	750	825	1400	950	1025	500	575	650	1600	1150	1225	700	775	850	250	325	400	475
B	816	952	646	697	1088	782	833	476	527	578	1224	918	969	612	663	714	306	357	408	459
C	1050	1330	700	805	1610	980	1085	350	455	560	1890	1260	1365	630	735	840	0	105	210	315
D	1056	1312	736	832	1568	992	1088	416	512	608	1824	1248	1344	672	768	864	96	192	288	384

Please make a choice:

A **B** **C** **D**

Figure A.2: Decision in the first round of a part.

Starting in round two of a part, the **results from the previous round** will be shown (see Figure A.3). In the first column, “Result,” you see the choices of all four players in the group. In the example figure it was “*B, D, B, C.*” The first letter (“*B*”) always represents your own choice whereas the following three letters (“*D, B, C*”) represent the choice of the other three members of your group. The position of a letter (choice of a group member) is always assigned to a specific group member and stays the same during a given part. In the example the “left” player chose *D*, the “middle” player chose *B*, and the “right” player chose *C*.

Round 2 of 51																						
	AAA	AAB	AAC	AAD	ABB	ABC	ABD	ACC	ACD	ADD	BBB	BBC	BBD	BCC	B C D	BDD	CCC	CCD	CDD	DDD		
A	1000	1200	750	825	1400	950	1025	500	575	650	1600	1150	1225	700	775	850	250	325	400	475		
B	816	952	646	697	1088	782	833	476	527	578	1224	918	969	612	663	714	306	357	408	459		
C	1050	1330	700	805	1610	980	1085	350	455	560	1890	1260	1365	630	735	840	0	105	210	315		
D	1056	1312	736	832	1568	992	1088	416	512	608	1824	1248	1344	672	768	864	96	192	288	384		
Results from the previous round																						
Results				Your Choice and Points				Choice with the most Points														
Choice of all Players: B, D, B, C				Choice: B Points: 663								Choice: C Points: 1365										
Please make a choice:										A		B		C		D						

Figure A.3: *BestOnly treatment*: New choice and the result of the previous round starting from the second round.

Round 2 of 51																				
	AAA	AAB	AAC	AAD	ABB	ABC	ABD	ACC	ACD	ADD	BBB	BBC	BBD	BCC	B C D	BDD	CCC	CCD	CDD	DDD
A	1000	1200	750	825	1400	950	1025	500	575	650	1600	1150	1225	700	775	850	250	325	400	475
B	816	952	646	697	1088	782	833	476	527	578	1224	918	969	612	663	714	306	357	408	459
C	1050	1330	700	805	1610	980	1085	350	455	560	1890	1260	1365	630	735	840	0	105	210	315
D	1056	1312	736	832	1568	992	1088	416	512	608	1824	1248	1344	672	768	864	96	192	288	384
Results from the previous round																				
Results				Your Choice and Points				Choice and Points of the Group Members												
Choice of all Players: B, D, B, C				Choice: B Points: 663				Choice: D Points: 1248				Choice: B Points: 663				Choice: C Points: 1365				
Please make a choice:																				
<div>A</div> <div>B</div> <div>C</div> <div>D</div>																				

Figure A.3: *FullInfo treatment*: New choice and the result of the previous round starting from the second round.

In the second column, “Your Choice and Points,” you will see your own choice and the points you earned in the previous round. In the example in Figure A.3 you can see in the payoff table that you earned 663 points because you chose *B* (row “*B*” in the table) and the other group members chose *D*, *B*, *C* (column “*B C D*”). In the table, the column representing the choice combination of the group members is highlighted in yellow. The columns are ordered alphabetically for a better overview.

[*Main Experiment BestOnly treatment:*] The last column, “Choice with the most Points,” shows the choice which earned the most points in the previous round and the points earned with that choice. This information is always highlighted in yellow.

[*Main Experiment FullInfo treatment, Replications:*] The last column, “Choice and Points of the Other Group Members,” shows the choices and how many points the other group members earned in the previous round. The ordering of the group members is the same as in the first column, “Result” (“left” player – *D*, “middle” player – *B*, and “right” player – *C*). The choice and points of the player who earned the most points in the previous round is highlighted in yellow. The column “Your Choice and Points” is also highlighted in yellow if you earned the most points in the previous round. In case of a tie the choice and points of multiple players will be highlighted.

[*Load replication:*] **Additional Points:** In every round you have the opportunity to earn additional points. To earn these points you have to **memorize** a number you see before you enter the decision phase. The number consists of **7 digits** and will be displayed for 10 seconds on the screen (see Figure A.4). Then the decision phase in which you have to choose among *A*, *B*, *C*, or *D* starts.

Round 2 of 51																				
	AAA	AAB	AAC	AAD	ABB	ABC	ABD	ACC	ACD	ADD	BBB	BBC	BBD	BCC	BCD	BDD	CCC	CCD	CDD	DDD
A	1000	1200	750	825	1400	950	1025	500	575	650	1600	1150	1225	700	775	850	250	325	400	475
B	816	952	646	697	1088	782	833	476	527	578	1224	918	969	612	663	714	306	357	408	459
C	1050	1330	700	805	1610	980	1085	350	455	560	1890	1260	1365	630	735	840	0	105	210	315
D	1056	1312	736	832	1568	992	1088	416	512	608	1824	1248	1344	672	768	864	96	192	288	384

Memorize this number:

8417253

Figure A.4: Example for a number consisting of 7 digits for the additional points.

After the decision phase you have to enter the whole number in the correct order (see Figure A.5). If you **correctly** enter the number you earn additionally to the points earned from the decision phase **750** points. You have to enter the number (without spaces) and click “OK” within 10 seconds otherwise it will be counted as false automatically. For a false input you will not receive any points. The additional points will be added to your points from the decision phase at the end of the experiment.

Do you have any questions? If so, please raise your hand and wait.

Round 2 of 51

	AAA	AAB	AAC	AAD	ABB	ABC	ABD	ACC	ACD	ADD	BBB	BBC	BBD	BCC	BCD	BDD	CCC	CCD	CDD	DDD
A	1000	1200	750	825	1400	950	1025	500	575	650	1600	1150	1225	700	775	850	250	325	400	475
B	816	952	646	697	1088	782	833	476	527	578	1224	918	969	612	663	714	306	357	408	459
C	1050	1330	700	805	1610	980	1085	350	455	560	1890	1260	1365	630	735	840	0	105	210	315
D	1056	1312	736	832	1568	992	1088	416	512	608	1824	1248	1344	672	768	864	96	192	288	384

Enter the number (without spaces):

OK

Figure A.5: Input of the number.

On your screen you will find some control questions for the experiment. Please answer these questions. If you have trouble answering the questions, please raise your hand. The experiment will start as soon as all participants answered the questions correctly.