Learning in a Black Box*

H.H. Nax[†] M.N. Burton-Chellew[‡] S.A. West[§] H.P. Young[¶]

First version: April 18, 2013
This version: February 18, 2014

*We thank Margherita Comola for help at an early stage of the project and Johannes Abeler, Guillaume Hollard, Richard Bradley, Fabrice Etilé, Philippe Jehiel, Thierry Verdier, Vince Crawford, Colin Camerer, Muriel Niederle, Edoardo Gallo, Amnon Rapoport, Guillaume Fréchette, Ozan Aksoy, Margaret Meyer, and participants of the Economics and Psychology seminar at MSE Paris, the Choice Group at LSE, the 24th Game Theory Workshop at Stony Brook, and the CESS seminar at Nuffiel College for helpful discussions, suggestions and comments. Correspondence to: H. H. Nax, ETH Zürich, Clausiusstrasse 50-C3, 8092 Zurich, Switzerland.

†Heinrich H. Nax, Department of Social Sciences-SOMS, ETH Zurich. Nax's research was in part supported by the Office of Naval Research grant N00014-09-1-0751 and under project NET of the Agence Nationale de Recherche.

‡Maxwell N. Burton-Chellew, Nuffield College, Oxford & Department of Zoology, University of Oxford (The Tinbergen Building S Parks Rd, Oxford OX1 3PS, United Kingdom). Burton-Chellew's research was supported by the European Research Council.

§Stuart A. West, Department of Zoology, University of Oxford (The Tinbergen Building S Parks Rd, Oxford OX1 3PS, United Kingdom). West's research was supported by the European Research Council.

¶H. Peyton Young, Department of Economics, University of Oxford (Nuffield College, New Rd, Oxford OX1 1NF, United Kingdom). Young's research was supported in part by the Office of Naval Research grant N00014-09-1-0751.

Abstract

Many interactive environments can be represented as games, but they are so large and complex that individual players are mostly in the dark about others' actions and the payoff structure. This paper analyzes learning behavior in such 'black box' environments, where players' only source of information is their own history of actions taken and payoffs received. The context of our analysis are decisions in voluntary contributions games with high and low rates of return. We identify four robust features of the players' learning dynamics: search volatility, search breadth, inertia, and directional bias. These features are present when players have no information about the game, and also when they have full information.

JEL classifications: C70, C73, C91, D83, H41

Keywords: black box, learning, information, public goods game

1 Introduction

Some games are so complex and involve so many individuals that for all practical purposes the game itself is unknown to the players themselves. Examples include bidding in online markets, threading one's way through urban traffic, or participating in a group effort where the actions of the other members of the group are difficult to observe (guerrilla warfare, neighborhood watch programs, tax evasion). In each of these cases a given individual will have only the haziest idea of what the other players are doing and what their payoffs are, but their own payoffs are strongly influenced by the actions of the others. How do individuals learn in such environments and under what circumstances does their learning behavior lead to equilibrium?

In this paper we investigate these questions in a laboratory environment. Players take actions and receive payoffs that depend on others' actions, but in the baseline case they have no information about the others and they do not know what the overall structure of the game is. This complete lack of structural information distinguishes our experimental set-up from other 'low-information' settings such as Rapoport et al. (2002) and Weber (2003), who provide information about the structure of the game but withhold information about the outcome of play and the distribution of players' types. This also distinguishes our experimental design from Friedman et al. (2012), who withhold information about the payoff structure of the game but do provide information about other players' actions and payoffs, or from Oechssler and Schipper (2003), where players learn the payoff structure on the basis of their own payoffs and information about other's actions.[1]

The underlying games played in our experiments have the structure of 'voluntary contributions games'. Rates of return on contributions are either low or high, respectively resulting either in the standard 'public goods game' where free-riding is the dominant strategy or in a contribution game without social dilemma where fully contributing is the dominant strategy. Subjects are faced with repeated decisions to allocate their monetary endowment between two accounts, but do not know that they play a contributions game at all. Our experiment ties in with a rich literature on public goods experiments (for surveys of this literature see Ledyard, 1995 and Chaudhuri, 2011), but unlike previous studies our focus is entirely different. In contrast to the existing literature, our aim is not to disentangle learning from social preferences (e.g. Andreoni, 1993, 1995; Palfrey and Prisbey, 1996, 1997; Goeree et al., 2002; Ferraro and Vossler, 2010; Fischbacher and Gachter, 2010; Burton-Chellew and West, 2013).[2] Rather, we study the behavioral adjustment regularities that govern how agents learn in a minimum-information environment.[3]

The closest set-up to ours is that of Bayer et al. (2013) who also reveal no information about the structure of the game or the actions and payoffs of the other players in the context of public goods games with low rates of return. Their framework also introduces an element of 'confusion' in which players are told that the game structure (and resulting payoffs) may change over the course of the experiment. In our setting, by contrast, the structure of the game and the payoffs, although unknown to the agents, remain explicitly fixed and the agents know that. Another early antecedent of our experiment are the repeated two-by-two zero sum games used to test Markov learning models (Suppes and Atkinson, 1959). However, subjects in their setting had only two available actions and it was permissible at that time to explicitly 'fool' subjects into thinking that they were not even playing a game.[4]

In this paper, we study voluntary contributions games in a 'black box' environment. As

[1]See Erev and Haruvy (2013) for a recent survey of equilibrium learning in this spirit. See also Ben Zion et al. (2010) for a low-information portfolio allocation experiment.

[2]In fact, social preference motivations (e.g. Fehr and Schmidt, 1999; Fehr and Gächter, 2000, 2002; Fehr and Camerer, 2007) cannot drive subjects' behavior in our baseline case because all information about the structure of the game and about others is withheld.

[3]We also complement previous studies of experience effects in voluntary contributions games (e.g. Marwell and Ames, 1981; Isaac et al., 1984; Isaac and Walker, 1988) by analysis of experience effects in a minimum-information setting.

[4]The observed learning in Suppes and Atkinson (1959) follows reinforcement (Erev and Roth, 1998; Erev and Rapoport, 1998).

the game proceeds, the subjects know only the amount that they themselves contributed and the payoff that they received as a result. Learning in such an environment is said to be *completely uncoupled* (Foster and Young, 2006; Hart and Mas-Colell, 2003, 2006; Young, 2009). In this setting, many of the standard learning rules used in the empirical literature, such as experience-weighted attraction, k-level reasoning and imitation, do not apply because these rules use the actions of the other players as inputs (see Björnerstedt and Weibull, 1993; Stahl and Wilson, 1995; Nagel, 1995; Ho et al., 1998; Camerer and Ho, 1999; Costa-Gomes et al., 2001; Costa-Gomes and Crawford, 2006; Crawford, 2013).

Nevertheless our experiments show that learning 'inside the box' can and does take place at rates that are comparable to previous studies (see Ledyard, 1995 and Chaudhuri, 2011). Moreover, the learning process exhibits certain distinctive features:

(*i*) *Search volatility.* A decrease in a player's realized payoff triggers greater variance in his choice of action next period, whereas an increase in his realized payoff results in lower variance next period.

(*ii*) *Search breadth.* The absolute value of a player's adjustment is on average larger after a payoff decrease than after a payoff increase.

(*iii*) *Inertia.* A player is more likely to stick with his previous contribution after a payoff increase than after a payoff decrease.

(*iv*) *Directional bias.* If an increase resulted in success this period, the player's contribution next period will tend to be higher than if it resulted in failure. Similarly, if a decrease resulted in success, the player's contribution next period will tend to be lower than if it resulted in failure.

These four learning components have antecedents in various strands of the learning literature, although they have not been given the explicit formulation that we do here. In particular, search volatility, search breadth, and inertia have been proposed in various forms in biology (Thuijsman et al., 1995), organizational learning (March, 1991), computer science (Eiben and Schippers, 1998), and machine learning (Nowak and Sigmund, 1993). Learning rules with related properties are variously known as 'exploration versus exploitation', 'win stay, lose shift' or 'fast and slow' learning. The basic idea is that an agent tends to keep playing a strategy that leads to high payoffs, and switches when payoffs decrease.[5] Financial market traders have also been shown to exhibit this kind of behavior (Coates, 2012).

[5]The first names given to behavioral rules based on this principle were 'simpleton' (Rapoport and Chammah, 1965) and 'Pavlov' (Nowak and Sigmund, 1993).

The closest antecedents in the literature are in biology, and in particular, models of foraging behavior of bees (Thuijsman et al., 1995; Motro and Shmida, 1995). Indeed, the model of Thuijsman et al. (1995) exhibits three of the four learning components of our model: search volatility, search breadth, and inertia. Bees typically forage for nectar by moving from one flower to another within a given 'patch'. However, if several attempts lead to flowers with little or no nectar, the bee will fly to a new randomly chosen patch that lies far away from the previous one. This behavior is called 'near-far' learning (Thuijsman et al., 1995, Motro and Shmida, 1995). In other words, both the breadth and volatility of search increase after (repeated) failures. Furthermore, behavior is inertial in that it follows the principle that the bee will "stay in the patch as long as you find food and leave otherwise".[6] Closely related learning models have recently been proposed in game theory; in particular there is a family of 'trial-and-error' learning rules that leads to Nash equilibrium in generic n-person games (Foster and Young, 2006; Germano and Lugosi, 2007; Marden et al., 2009; Young, 2009; Pradelski and Young, 2012).

An additional feature of our framework is the one-dimensional nature of the strategy space, which induces directional bias in the agents' adjustments. If a higher contribution resulted in a higher payoff this period, a player will tend to make a larger contribution next period than if his previous contribution resulted in a lower payoff.[7] Hence, different regions of the strategy space are explored with higher probability given the directional feedback from previous adjustments. Directional bias has antecedents in the experimental literature on 'directional learning' and 'aspiration adjustment theory' (Sauermann and Selten, 1962; Cross, 1983; Selten and Stoecker, 1986; Selten and Buchta, 1998).[8] A theoretical model of directional learning that leads to Nash equilibrium in two-by-two games with one-dimensional strategies has recently been proposed by Laslier and Walliser (2011). Bayer et al. (2013) investigate another directional learning model in the context of voluntary contributions games. Our results confirm the presence of such directional biases, as we shall show in section 3.

In this paper we show that our learning model helps explain the behavior of subjects who are learning in a 'black box' environment. The contribution of this paper is to identify and to test the four key features mentioned above when subjects have no information about the game and they cannot observe others' behavior. Our framework is therefore fundamentally different from most other models in the literature on experimental games,

[6]See Thuijsman et al. (1995), p. 309.

[7]Similarly, if a *lower* contribution resulted in a higher payoff, he will tend to make a smaller contribution next period than if his prior contribution resulted in a lower payoff.

[8]See Harstad and Selten (2013) for a recent survey and experimental evidence from Tietz and Weber, 1972; Roth and Erev, 1995; Erev and Roth, 1998; Erev and Rapoport, 1998).

which generally assume observability of others' actions and 'some' knowledge of the game (examples include experience-weighted attraction, k-level reasoning, and imitation). As we shall see, all four features of our model are confirmed at high levels of statistical significance.

We conclude that even when subjects have virtually no information about the game, learning does take place and exhibits the four distinctive patterns identified above. Furthermore, these features are still present even when more information becomes available and/ or players gain experience. While more experience and more information in the different information treatments lead to different contribution levels and to faster convergence rates for the two rates of return compared to the black box treatment, the four components of our learning model are still present in the non-black box treatments.

The paper is structured as follows. In section 2 we describe the experimental set-up in detail. In section 3 we present the black box findings, which we compare with the other information treatments in section 4. We conclude in section 5. The appendix contains experimental instructions and supplementary regression tables.[9]

2 Experimental set-up

A total of 236 subjects, in 16 separate sessions involving 12 or 16 subjects, participated in our public goods experiments yielding a total of 18,880 observations. Participants were recruited from a subject pool that had not previously been involved in public goods experiments.[10] The subjects were not limited to university students, but included different age groups with diverse educational backgrounds.[11] In the experiment, a group of subjects plays several voluntary contributions games, where each game was repeated for twenty rounds with randomly allocated subgroups in each round. Games differ with respect to two *rates of return* ('low' and 'high') such that either 'free-riding' or 'fully contributing' is the dominant strategy. These are played under three different information treatments ('black box', 'standard' and 'enhanced'). In this section, we shall describe the underlying stage game, the structure of each repeated game, and the different information treatments.

[9] See also Supporting Information for Burton-Chellew and West 2013.

[10] The subjects were recruited through ORSEE (Online Recruiting System for Economic Experiments; Greiner, 2004) and subjects who listed prior participation in public goods games were excluded.

[11] The experiment was programmed and conducted with the software z-Tree (Fischbacher, 2007). All experiments were conducted at the Centre for Experimental Social Sciences (CESS) at Nuffield College, University of Oxford.

2.1 Linear public goods game

Consider the following linear public goods game, known as the "voluntary contributions game". Every player i in population $N = \{1, 2, ..., n\}$ makes a nonnegative, real-numbered contribution, c_i, from a finite budget $B > 0$. Given a resulting vector of all players' contributions, $\mathbf{c} = \{c_1, c_2, ..., c_n\}$, for some *rate of return*, $e \geq 1$, the public good is provided in the amount $R(\mathbf{c}) = e \sum_{i \in N} c_i$, and split equally amongst the players.[12] Given others' contributions $\mathbf{c}_{-i}$, player i's contribution c_i results in the payoff

$$\phi_i = \frac{e}{n} \sum_{i \in N} c_i + (B - c_i),$$

where $\frac{e}{n}$ is the marginal rate of return. Write $\boldsymbol{\phi}$ for the payoff vector $\{\phi_1, ..., \phi_n\}$.

Nash equilibria. Depending on whether the rate of return is *low* ($e < n$) or *high* ($e > n$), an individual contribution of zero ('free-riding') or B ('fully contributing') is the strictly dominant strategy for all players. The respective Nash equilibrium results in either nonprovision or full provision of the public good.

2.2 Repeated game

In each experimental session, the same population S (with $|S| = 12$ or 16) plays four 'phases' where each phase is a separate twenty-times repeated voluntary contributions game. In each period t within a given phase, players in S are randomly matched in groups of four to play the voluntary contributions game.[13] At the start of each period, each subject is given a new budget $B = 40$, of which he can invest any amount; however, subjects cannot invest money carried over from previous rounds. The rate of return is either *low* (set at $e = 1.6$) or *high* (set at $e = 6.4$) throughout a given phase.[14] Write N_4^t for any of the four-player groups matched at time t, and ρ_4^t for the partition of S into such groups. Given others' contributions $\mathbf{c}_{-i}^t$, each i receives a total of

$$\phi_i = \sum_{t=1,...,20} \phi_i^t = \sum_{t=1,...,20} \left(\frac{e}{4} \sum_{j \in N_4^t} c_j^t + (40 - c_i^t) \right). \tag{1}$$

[12] If $e < 1$, $R(\mathbf{c})$ is a public bad.

[13] This follows random 'stranger' rematching of Andreoni (1988).

[14] The Nash equilibrium payoff of the stage game is 40 when $e = 1.6$, and 256 when $e = 6.4$.

For every i, ϕ_i represents a real monetary value that is paid after the game.[15] Note that, due to random rematching of players into groups throughout the experiment, for every player i, his relevant group N_4^t in the above expression 1 (i.e. such that $i \in N_4^t$) typically is a different one in each period t.

2.3 Information treatments

Each experimental session (involving four phases) is divided into two 'stages': phases one and two of the session constitute stage one, phases three and four constitute stage two. Each phase is a twenty-round voluntary contributions game with either the low ($e = 1.6$) or the high ($e = 6.4$) rate of return, and in each stage both rates of return are played. The two stages differ with respect to the information treatment.

Before the experiment begins, subjects are told that two separate experiments will be conducted, each stage consisting of two games. At no point before, during, or in between the two separate stages of the experiment are players allowed to communicate. Depending on which treatment is played, the following information is revealed at the start of each stage.[16]

All treatments. During each phase, the same underlying game is repeated for 20 periods. Each player receives 40 monetary units each period, of which he can invest any amount. After investments are made, each player earns a nonnegative return each period which, at the end of each game, he receives together with his uninvested money according to a known exchange rate into real money.

Black box. No information about the structure of the games or about other players' actions or payoffs is revealed. Subjects play two voluntary contributions games (one with the high and one with the low rate of return). As play proceeds, subjects only know their own contributions and payoffs, but do not learn those of others.

Standard. The rules of the game are revealed, including production of the public good, high and low rate of return, and how groups form each period. As the game proceeds, players receive a summary of the relevant contributions in their group at the end of each period.[17]

[15] One hundred coins are worth £0.15. The maximal earnings over the whole experiment therefore amount to £19.2.

[16] See Appendix A for the full instructions and Appendix B for the output screens displayed during the experiment for each treatment.

[17] This follows the standard information treatment design as in Fehr and Gachter (2002) and random 'stranger' rematching of subjects as introduced by Andreoni (1988).

Enhanced. In addition to the information available in the standard treatment, the payoffs of the other group members are explicitly calculated, and players receive a summary of their own and other players' payoffs in their group at the end of each period.

In each experimental session, every player plays two black box games (one with the high and one with the low rate of return), and either two standard or two enhanced information games (again one high and one low). Either the black box treatment occurs first, or a non-black box treatment occurs. Of the total of 18,880 observations, 9,440 are black box, 4,640 are standard, and 4,800 are enhanced. The order of high and low rates of return and of treatments is different in each session. Sessions lasted between fifty and sixty minutes, and subjects earned between £6.20 and £15.50 (mean earnings were £12.40).

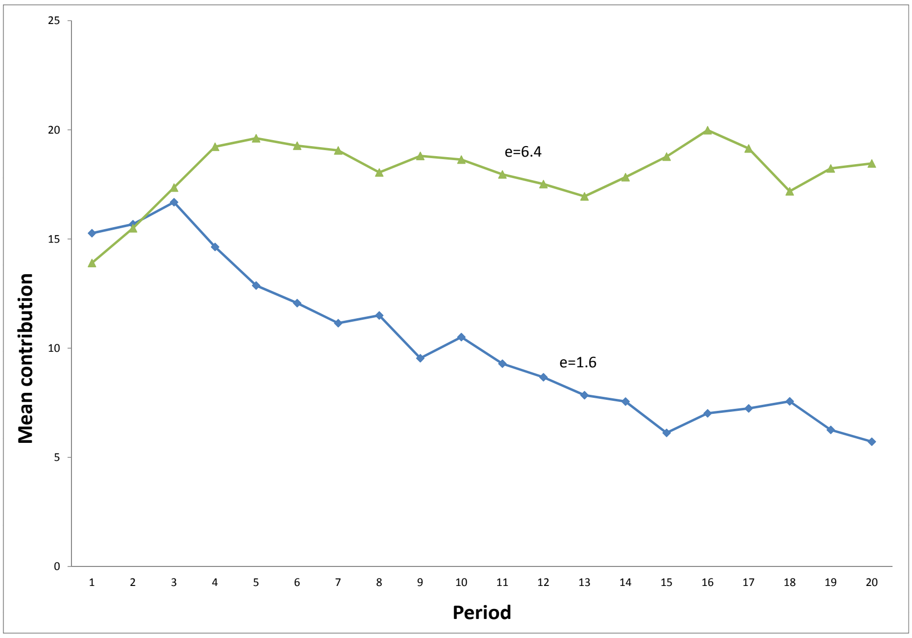
Black box versus grey box

Black box is our main treatment, and the analysis is based on sessions when black box is played first. Recall that we require the recruiting system (ORSEE) to select only 'first timers' of public goods experiments. Thus, when black box is played first, subjects are not likely to have prior knowledge of the structure of the game.

'Black box' versus 'grey box'. *We shall reserve the term 'black box' for the case when the no information treatment is played first. Sessions when it is played after the standard or enhanced treatments will be called 'grey box'.*

In grey box, subjects are told that a new and separate experiment will be conducted and that all information except for their own payoffs will be withheld. Since play in these sessions was preceded by two voluntary contributions games where subjects received full information about the structure of the game, they might (or might not) think that the grey box treatment has a similar structure. However, they will still be unable to infer others' contributions and the underlying rates of return from the information they receive. We make this distinction between black box and grey box because our black box learning model applies most evidently in the (pure) black box environment where more sophisticated learning models cannot apply because of the complete absence of information (about the structure and about others' actions).

Figure 1: Black box play.



In the free-riding game ($e = 1.6$), contributions deteriorate at a faster rate than contributions rise in the full contribution game ($e = 6.4$).

3 Black box learning

In this section, we present and test our learning model based on the analysis of the pure black box data, that is, those sessions when black box was played before the standard or enhanced treatments. There were eight sessions and 4,960 observations. Recall that subjects in the black box have no knowledge of the structure of the game and receive no information about others' actions or payoffs as the game goes on. Moreover, subjects are recruited to be 'first timers', that is, they have no prior experience playing public goods games. In the subsequent analysis, unless stated otherwise, "significance" refers to a 95 percent confidence ($p < 0.05$) with which the null can be rejected and to a 90 percent confidence interval ($p > 0.1$) when the null cannot be rejected.

Figure 1 illustrates black box play, averaged over individuals and sessions. The contribution patterns associated with the low rate of return are comparable with those from previous studies; for a recent review of this literature see Chaudhuri (2011). The high rate of return has been less frequently studied. For a comparison of the convergence rates (for black box and for the comparison with the other information treatments), see

Burton-Chellew and West (2013).

3.1 The learning model

We begin by establishing some terminology.

Adjustment. The *adjustment* of player i in period $t+1$ is $c_i^{t+1} - c_i^t$. The adjustment is an *increase* if $c_i^{t+1} > c_i^t$, a *decrease* if $c_i^{t+1} < c_i^t$, and a *zero adjustment* if $c_i^{t+1} = c_i^t$.

Breadth. The *breadth* of an adjustment by player i in period $t+1$ is the absolute value of the adjustment, $|c_i^{t+1} - c_i^t|$.

Success versus failure. Player i experiences *success* in period $t+1$ if his realized payoff does not go down ($\phi_i^{t+1} \geq \phi_i^t$); otherwise he experiences *failure.*

Bias. The *bias* in adjustments of player i in period $t+1$ is the difference between the expected adjustment after success and the expected adjustment after failure, that is, $\mathbf{E}(c_i^{t+1} - c_i^t | \phi_i^t \geq \phi_i^{t-1}) - \mathbf{E}(c_i^{t+1} - c_i^t | \phi_i^t < \phi_i^{t-1})$.

Our model, called **SEARCH**, consists of the following four components:

(i) *Search volatility.* Failure triggers a greater variance in non-zero adjustments than does success.

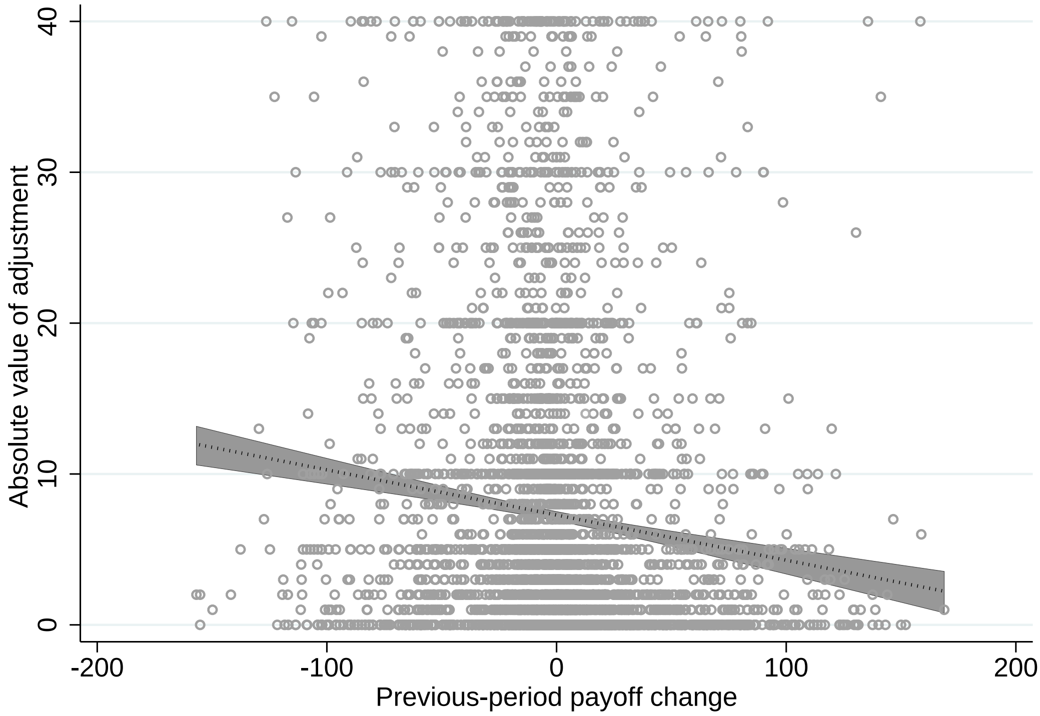
(ii) *Search breadth.* The absolute size of non-zero adjustments is larger after failure than after success.

(iii) *Inertia.* Zero adjustments are less likely after failure than after success.

(iv) *Directional bias.* If an increase in contribution resulted in success this period, the player's contribution next period will tend to be higher than if the prior contribution resulted in failure. Similarly, if a decrease in contribution resulted in success, the player's contribution next period will tend to be lower than if the prior contribution resulted in failure.

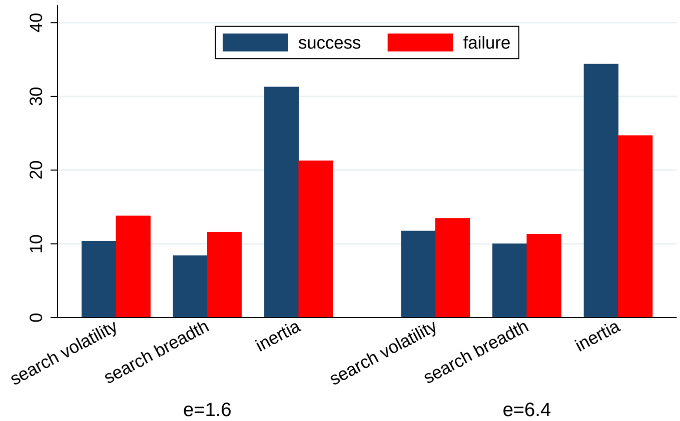
Consider the following example. Suppose an agent contributes 10 in one period and 20 in the next. Search volatility means that if 20 results in a higher realized payoff (success) to the player, then his next-period adjustment is drawn from some distribution with lower variance than if 20 resulted in a lower realized payoff (failure). Search breadth states that the absolute value of the adjustment, in case an adjustment is made, tends to be larger after failure than after success. Inertia states that he is more likely to stick with 20 again in case of success than in case of failure. Finally, directional bias implies that he will tend to contribute more if 20 was a success than if 20 was a failure.

Figure 2: Black box adjustments.



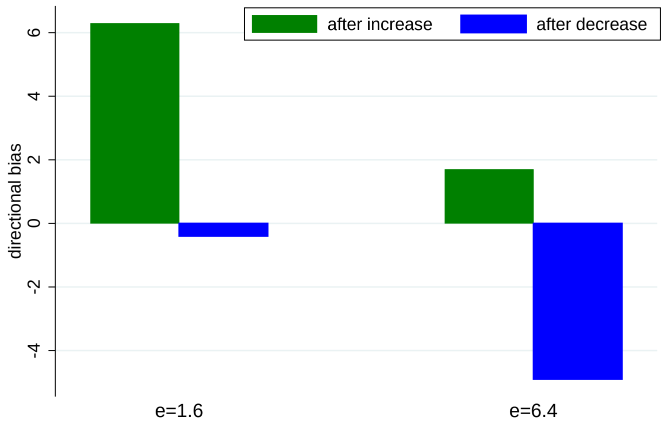
The figure shows absolute values of adjustments in the current period as a function of the last-period payoff change. The linear regression reveals that larger adjustments tend to lie on the side of previous-period failures, while smaller and inertial adjustments tend to lie on the success-side.

Figure 3: Search volatility, search breadth, and inertia.



The bar charts summarize search volatility, search breadth, and inertia conditional on success versus failure under black box information.

Figure 4: Directional bias.



The bar charts summarize the directional bias, that is, the difference between adjustments after success and adjustments after failure conditional on previous-period increase versus decrease under black box information.

We find that all of these features are present in the data. Consider Figure 2, which plots the absolute value of adjustments against the performance of previous contributions. The scatterplot suggests that adjustments after failure are more volatile; larger adjustments occur more often after failure than after success; and lower or inertial adjustments are more likely after success than after failure. A summary of all four SEARCH components is given in Figures 3 and 4 . These patterns are confirmed by statistical tests of significance, which are summarized in Tables 1 and 2. Comparing columns (c) and (d) in Table 1, we make the following observations: (i) search volatility breadth is higher after failure than after success; (ii) search breadth is higher after failure than after success; and (iii) inertia is lower after failure than after success. From columns (b) and (c) in Table 2, we see that (iv) directional bias is such that adjustments after increases (decreases) are higher (lower) in case of success than in case of failure.

The tests of significance were conducted as follows. For search volatility, we applied Levene's test, which is a nonparametric test for the equality of variances in different samples. The null hypothesis of equal variances following success and failure, which would imply no search volatility, is rejected with 99 percent confidence. This holds for both high and low rates of return and for different orders of the games (see Appendix C, Table 3, Tests 1-3).

Next, we test the hypothesis of equal search breadth conditional on success versus failure. We regress the absolute values of the non-zero adjustments conditional on success versus failure, controlling for phase, group, period and individual fixed effects with individual-

Table 1: Summary statistics for search volatility, search breadth, and inertia (black box treatment)

(a) search component	(b)	(c) failure	(d) success	(e) difference (c)-(d)
search volatility				
standard deviation of adjustment	standard deviation	13.7	11.1	2.6**
	# observations	2266	2198	
search breadth				
absolute value of non-zero adjustments	mean	11.4	9.2	2.2**
	# observations	1746	1476	
inertia				
probability of non-zero adjustment	relative frequency	0.23	0.33	0.10**
	# observations	520	722	

*: $p < 0.05$; **: $p < 0.01$.

level clustering. The average adjustment following success is smaller than after failure at a 1 percent level of significance (see Appendix C, Table 5). The breadth of search is significantly larger in periods 1-9 than in periods 10-20. Group, phase, and period fixed effects are not significant.

To test the hypothesis of equal inertia after success and failure, we use an ordered probit regression of the inertia rate controlling for phase, group, period and individual fixed effects with individual-level clustering. Success turns out to have a significantly positive effect on the inertia rate (see Appendix C, Table 5). Moreover, there is significantly less inertia in the first phase than in the second phase. Period fixed effects are negative until period six, suggesting that inertia tends to increase over time. Group fixed effects are not significant.

Finally, we analyze the directional patterns of adjustments. Our hypothesis of directional bias states that (*i*) if an increase leads to success, the player's next-period contribution will tend to be higher than if it leads to failure; (*ii*) if a decrease leads to success, the player's contribution next period will tend to be lower than if it leads to failure. To test this

Table 2: Summary statistics for directional bias (black box treatment)

(a) mean non-zero adjustment	(b) success	(c) failure	(d) difference (b)-(c)
after *increase*	-1.6	-6.5	4.9**
after *decrease*	2.5	5.5	-3.0**

*: $p < 0.05$; **: $p < 0.01$.

hypothesis, we regress the difference between adjustments after success and adjustments after failure, controlling for the direction of the previous adjustment as well as for phase, group, period and individual fixed effects with individual-level clustering. The directional bias is confirmed, for both increases and decreases in prior-period contributions, at a high level of statistical significance ($p < 0.01$). The details are given in Appendix C, Table 6.

3.2 Summary

Our black box findings may be summarized as follows. First, there is strong evidence for SEARCH: search volatility and search breadth are both larger after failure than after success; inertia is larger after success than after failure; and there is directional bias, meaning that increases (decreases) tend to be followed by larger (smaller) contributions after success than after failure. Despite subjects' lack of information, however, the overall pattern of the dynamics is in line with previous experiments. In free-riding games ($e = 1.6$), for example, contributions deteriorate at almost the same rate as in previous studies (Ledyard, 1995; Chaudhuri, 2011; Burton-Chellew and West, 2013).

4 Non-black box data

In this section, we assess our black box findings in light of the data from grey box and from the other two treatments, standard and enhanced. In particular, we shall investigate whether SEARCH describes only black box behavior or whether its components also persist in the other settings, that is, where players gain experience and/ or have explicit information about the structure of the game and others' actions (and payoffs).

Figure 5: Mean play all treatments (black box, standard, enhanced).

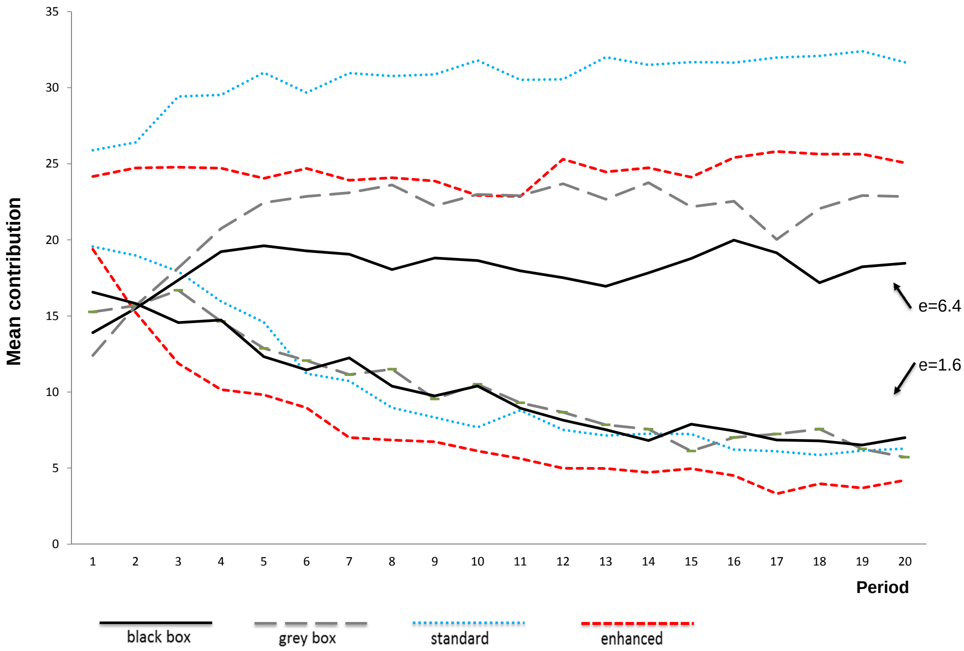


Figure 5 illustrates play in all treatments.

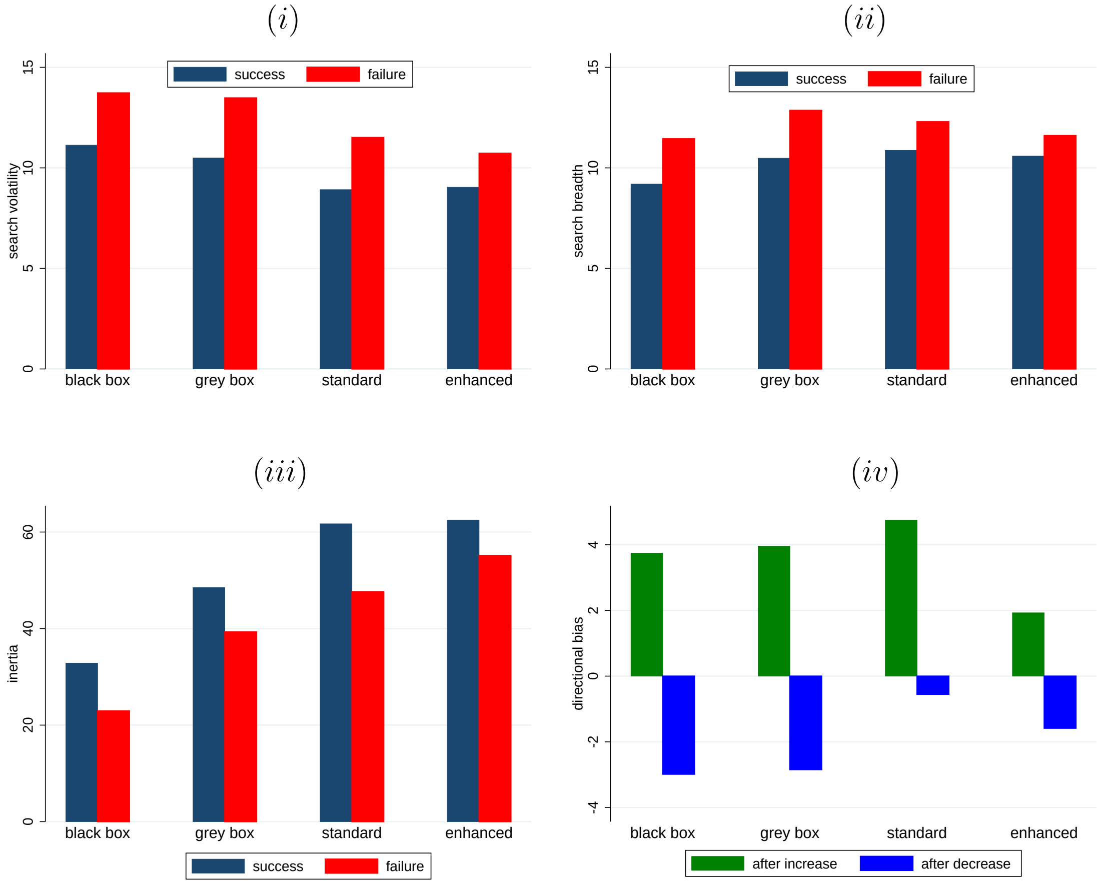
Figure 5 illustrates play in the different treatments, averaged over individuals and sessions. Different information clearly makes a difference in the level of contributions and/ or convergence rates (see Burton-Cellew and West, 2013). However, the features of SEARCH persist in all sessions and in all treatments, as is summarized in Figure 6.[18]

4.1 Black box versus grey box

First, we shall investigate 'grey box' play. Recall that grey box means that black box is played after either enhanced or standard. Even though subjects are explicitly told that a separate experiment is started after the first stage of the experiment, players may or may not make inferences about the game structure. We shall investigate the consequences of this effect in comparison with (pure) black box behavior. Our analysis reveals that although there are differences in average levels of contribution in the two cases (Figure 5), all of the features of SEARCH are robust for both rates of return. In particular, the levels of search volatility and search breadth are not significantly different in the two cases (Appendix C, Tables 7 and 8). The differences in inertia rates conditional on

[18]Appendix C, Figure 8 contains a full summary of all SEARCH components in all treatments for both rates of return separately.

Figure 6: SEARCH (all treatments).



The bar charts summarize the four search components in all treatments. The respective panels are: (i) search volatility, (ii) search breadth, (iii) inertia, (iv) directional bias.

success-versus-failure are similarly robust, however, the level of inertia is higher in grey box than in black box (Appendix C, Table 5). Directional bias is unchanged in terms of its size and significance levels compared to black box (Appendix C, Table 10). We conclude that grey box differs from black box with regard to the absolute level of inertia, but not qualitatively with respect to any of the SEARCH components.

4.2 Black box versus standard and enhanced

Finally, we consider whether the standard and enhanced treatments lead to different conclusions. The situation is summarized in Figure 6.[19] Search volatility, though smaller in absolute size in the standard treatment and even smaller in the enhanced treatment, is higher after failure than after success, just as in the black box case (see panel (i) in Figure 6). Moreover, this difference is significant at the 1 percent level (see Appendix C, Table

[19]See Appendix C, Figure 8 for a summary for both rates of return separately.

7). Search breadth exhibits the same success-versus-failure difference as in black box (see panel (*ii*) in Figure 6). It continues to be statistically significant at the 1 percent level (see Appendix C, Table 8). Panel (*iii*) in Figure 6 illustrates that inertia rates decrease after failure in all treatments. This success-failure differential in inertia rates is statistically significant in all four treatments, and there is higher absolute inertia in non-black box treatment data (see Appendix C, Table 5). Finally, panel (*iv*) in Figure 6 illustrates the presence of directional biases across all four treatments. We qualitatively confirm the increase-decrease directional bias from the black box data when analyzing the non-black box data at levels that are statistically significant (see Appendix C, Table 10).

5 Conclusion

Much of the prior empirical work on learning in games has focussed on situations where players have a substantial amount of information about the structure of the game and they can observe the behavior of others as the game proceeds. In this paper, by contrast, we have examined situations in which players have *no* information about the strategic environment. This takes us to the other end of the spectrum from *full* information. Players in such environments must feel their way to equilibrium based solely on the pattern of their own realized payoffs. Naturally, assuming that players would have *no* information whatsoever is somewhat extreme because players in practice will often have at least *some* information. To disentangle the role and to highlight the nature of adaptive learning versus best-reply dynamics, we implemented a black box environment in our experiment and compared the resulting behavior to play under intermediate and rich information treatments. We identified four key features of the resulting (completely uncoupled) learning dynamics – search volatility, search breadth, inertia, and directional bias. Although these components have precursors in both psychology and biology, they have not been given the precise formulation that we propose here, nor have they been subjected to rigorous testing in a laboratory environment. It turns out all four features are validated at high levels of statistical significance. Moreover, they are present even when players gain more experience and/ or have more information. Whether this remains true for other classes of games is an open question for future research.

References

J. Andreoni, "Why Free Ride? Strategies and Learning in Public Goods Experiments", *Journal of Public Economics 37*, 291-304, 1988.

J. Andreoni, "An experimental test of the public goods crowding-out hypothesis", *The American Economic Review 83*, 1317-1327, 1993.

J. Andreoni, "Cooperation in public-goods experiments: kindness or confusion?" *The American Economic Review 85*, 891-904, 1995.

R.-C. Bayer, E. Renner and R. Sausgruber, "Confusion and learning in the voluntary contributions game", *Experimental Economics 16,* 478-496, 2013.

U. Ben Zion, I. Erev, E. Haruvy, and T. Shavit, "Adaptive behavior leads to under-diversification", *Journal of Economic Psychology 31*, 985-995, 2010.

J. Björnerstedt, and J. Weibull, "Nash Equilibrium and Evolution by Imitation", in Arrow, K. and Colombatto, E. (Eds.), *Rationality in Economics*, New York: Macmillan, pp. 151-171, 1993.

M. N. Burton-Chellew and S. A. West, "Pro-social preferences do not explain human cooperation in public-goods games", *Proceedings of the National Academy of Science 110*, 216-221, 2013.

C. F. Camerer and T.-H. Ho, "Experience-weighted Attraction Learning in Normal Form Games", *Econometrica 67*, 827-874, 1999.

A. Chaudhuri, "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature", *Experimental Economics 14*, 47-83, 2011.

J. Coates, "The Hour Between Dog and Wolf. Risk Taking, Gut Feelings, and the Biology of Boom and Bust", *Penguin Press USA*, 2012.

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

M. Costa-Gomes and V. Crawford, "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study", *American Economic Review 96*, 1737-1768, 2006.

V. P. Crawford, "Boundedly Rational versus Optimization-Based Models of Strategic Thinking and Learning in Games", *Journal of Economic Literature 51*, 512-527, 2013.

R. Croson, "Partners and Strangers Revisited", *Economics Letters 53*, 25-32, 1996.

J. G. Cross, *A Theory of Adaptive Economic Behavior,* Cambridge University Press, 1983.

A. E. Eiben and C. Schippers, "On evolutionary exploration and exploitation", *Fundamenta Informaticae 35*, 35-50, 1998.

I. Erev and A. Rapoport, "Coordination, 'magic,' and reinforcement learning in a market entry game" *Games and Economic Behavior 23*, 146-175, 1998.

I. Erev and E. Haruvy, "Learning and the economics of small decisions", J. H. Kagel and A. E. Roth (Eds.), *The Handbook of Experimental Economics*, forthcoming, Princeton University Press, 2013.

I. Erev and A. E. Roth, "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria", *American Economic Review 88,* 848-881, 1998.

E. Fehr and S. Gächter, "Cooperation and Punishment in public goods experiments", *American Economic Review 90*, 980-994, 2000.

E. Fehr and S. Gächter, "Altruistic punishment in humans", *Nature 415*, 137-140, 2002.

E. Fehr and K. M. Schmidt, "A Theory of Fairness, Competition and Cooperation", *Quarterly Journal of Economics 114*, 817-868, 1999.

E. Fehr and C. Camerer, "Social neuroeconomics: the neural circuitry of social preferences", *Trends in Cognitive Sciences 11*, 419-427, 2007.

P. J. Ferraro and C. A. Vossler, "The source and significance of confusion in public goods experiments", *The B.E. Journal in Economic Analysis and Policy 10*, 53, 2010.

U. Fischbacher, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments", *Experimental Economics 10*, 171-178, 2007.

U. Fischbacher, and S. Gächter, "Social preferences, beliefs, and the dynamics of free riding in public good experiments", *The American Economic Review 100*, 541-556, 2010.

D. Foster and H. P. Young, "Regret testing: Learning to play Nash equilibrium without knowing you have an opponent", *Theoretical Economics 1*, 341-367, 2006.

D. Friedman, S. Huck, R. Oprea and S. Weidenholzer, "From imitation to collusion: Long-run learning in a low-information environment", *Discussion Papers in Economics of Change SP II 2012-301*, WZB , 2012.

F. Germano and G. Lugosi, "Global Nash convergence of Foster and Young's regret testing", *Games and Economic Behavior 60*, 135-154, 2007.

J. K. Goeree, C. A. Holt and S. K. Laury, "Private costs and public benefits: Unraveling the effects of altruism and noisy behavior" *Journal of Public Economics 83*, 255-276, 2002.

B. Greiner, "An Online Recruitment System for Economic Experiments", in: K. Kremer and V. Macho (Eds.): *Forschung und wissenschaftliches Rechnen 2003.* GWDG Bericht 63, 79-93, 2004.

R. M. Harstad and R. Selten, "Bounded-Rationality Models: Tasks to Become Intellectually Competitive", *Journal of Economic Literature 51*, 496-511, 2013.

S. Hart and A. Mas-Colell, "Uncoupled Dynamics Do Not Lead to Nash Equilibrium", *American Economic Review 93*, 1830-1836, 2003.

S. Hart and A. Mas-Colell, "Stochastic Uncoupled Dynamics and Nash Equilibrium," *Games and Economic Behavior 57*, 286-303, 2006.

T.-H. Ho, C. Camerer and K. Weigelt, "Iterated Dominance and Iterated Best-response in p-Beauty Contests", *American Economic Review 88*, 947-969, 1998.

M. Isaac, J. Walker and S. Thomas, "Divergent Expectations on Free Riding: An Experimental Examination of Possible Explanations", *Public Choice 43*, 113-149, 1984.

M. Isaac and J. Walker, "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism", *Quarterly Journal of Economics 103*, 179-199, 1988.

J.-F. Laslier and B. Walliser, "Stubborn Learning", *Ecole Polytechnique WP-2011-12*, 2011.

J. O. Ledyard, "Public Goods: A Survey of Experimental Research", in J. H. Kagel and A. E. Roth (Eds.), *Handbook of experimental economics,* Princeton University Press, 111-194, 1995.

J. G. March, "Exploration and Exploitation in Organizational Learning", *Organization Science 2*, 71-87, 1991.

J. R. Marden, H. P. Young, G. Arslan, J. Shamma, "Payoff-based dynamics for multiplayer weakly acyclic games", *SIAM Journal on Control and Optimization 48,* special issue on "Control and Optimization in Cooperative Networks", 373-396, 2009.

G. Marwell and R. Ames, "Economists Free Ride, Does Anyone Else?", *Journal of Public Economics 15*, 295-310, 1981.

U. Motro and A. Shmida, “Near-far search: an evolutionary stable foraging strategy”, *Journal of Theoretical Biology 173*, 15-22, 1995.

R. Nagel, “Unraveling in Guessing Games: An Experimental Study”, *American Economic Review 85*, 1313-1326, 1995.

M. Nowak and K. Sigmund, “A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner’s Dilemma game”, *Nature 364*, 56-58, 1993.

J. Oechssler and B. Schipper, “Can you guess the game you are playing?”, *Games and Economic Behavior 43*, 137-152, 2003.

T. R. Palfrey and J. E. Prisbrey, “Altruism, reputation and noise in linear public goods experiments”, *Journal of Public Economics 61,* 409-427, 1996.

T. R. Palfrey and J. E. Prisbrey, “Anomalous behavior in public goods experiments: how much and why?”, *The American Economic Review 87,* 829-846, 1997.

B. S. R. Pradelski and H. P. Young, “Learning Efficient Nash Equilibria in Distributed Systems”, *Games and Economic Behavior 75*, 882-897, 2012.

A. Rapoport and A. M. Chammah, *Prisoner’s Dilemma: A Study in Conflict and Cooperation*, University of Michigan Press, 1965.

A. Rapoport, D. A. Seale, and J. E. Parco, “Coordination in the aggregate without common knowledge or outcome information”, in: R. Zwick and A. Rapoport (Eds.): *Experimental Business Research*, pp. 69-99, 2002.

A. E. Roth and I. Erev, “Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term”, *Games and Economics Behavior 8,* 164-212, 1995.

H. Sauermann and R. Selten, “Anspruchsanpassungstheorie der Unternehmung”, *Zeitschrift für die Gesamte Staatswissenschaft 118*, 577-597, 1962.

R. Selten and J. Buchta, “Experimental Sealed Bid First Price Auction with Directly Observed Bid Functions”, in: Budescu, D., Erev, I., Zwick, R. (Eds.), *Games and Human Behavior,* Essays in Honor of Amnon Rapoport, pp. 79-104, 1998.

R. Selten and R. Stoecker, “End Behavior in Sequences of Finite Prisoner’s Dilemma Supergames: A Learning Theory Approach”, *Journal of Economic Behavior and Organization 7*, 47-70, 1986.

D. Stahl and P. Wilson, “On Players’ Models of Other Players: Theory and Experimental Evidence”, *Games and Economic Behavior 10*, 218-254, 1995.

P. Suppes and A. R. Atkinson, *Markov Learning Models for Multiperson Situations*, Stanford University Press, 1959.

F. Thuijsman, B. Peleg, M. Amitai and A. Shmida, "Automata, matching and foraging behavior of bees", *Journal of Theoretical Biology 175*, 305-316, 1995.

R. Tietz and H. Weber, "On the nature of the bargaining process in the Kresko-game", in: Sauermann, H. (Ed.), *Contributions to experimental economics*, Vol. 3, pp. 305-334, 1972.

R. A. Weber, "Learning' with no feedback in a competitive guessing game", *Games and Economic Behavior 44*, 134-144, 2003.

H. P. Young, "Learning by trial and error", *Games and Economic Behavior 65*, 626-643, 2009.

Appendices

Appendix A: Instructions

Participants received the following on-screen instructions (in z-Tree) at the start of the game. The set of instructions in standard and enhanced were the same, different instructions were given in black box. In black box, participants had to click an on-screen button saying, “I confirm I understand the instructions” before the game would begin. The same black box instructions were used for both rates of return. The standard/ enhanced instructions differ with respect to the relevant numbers for the two rates of return, and the example is adequately modified.

Black box

The following instructions were used in black box.

Beginning of instruction.

Instructions

Welcome to the experiment. You have been given 40 virtual coins. Each ‘coin’ is worth real money. You are going to make a decision regarding the investment of these ‘coins’. This decision may increase or decrease the number of ‘coins’ you have. The more ‘coins’ you have at the end of the experiment, the more money you will receive at the end.

During the experiment we shall not speak of £Pounds or Pence but rather of “Coins”. During the experiment your entire earnings will be calculated in Coins. At the end of the experiment the total amount of Coins you have earned will be converted to Pence at the following rate: 100 Coins = 15 Pence. In total, each person today will be given 3,200 coins (£4.80) with which to make decisions over 2 economic experiments and their final totals, which may go up or down, will depend on these decisions.

The Decision

You can choose to keep your coins (in which case they will be ‘banked’ into your private account, which you will receive at the end of the experiment), or you can choose to put some or all of them into a ‘**black box**’.

This ‘**black box**’ performs a mathematical function that converts the number of coins inputted into a number of coins to be outputted. The function contains a random component, so if two people were to put the same amount of coins into the ‘**black box**’, they would not necessarily get the same output. The number outputted may be more or less than the number you put in, but it will never be a negative number, so the lowest outcome possible is to get 0 (zero) back. If you chose to input 0 (zero) coins, you may still get some back from the box.

Any coins outputted will also be 'banked' and go into your private account. So, your final income will be the initial 40 coins, minus any you put into the '**black box**', plus all the coins you get back from the '**black box**'.

You will play this game 20 times. Each time you will be given a new set of 40 coins to use. Each game is separate but the '**black box**' remains the same. This means you cannot play with money gained from previous turns, and the maximum you can ever put into the '**black box**' will be 40 coins. And you will never run out of money to play with as you get a new set of coins for each go. The mathematical function will not change over time, so it is the same for all 20 turns. However as the function contains a random component, the output is not guaranteed to stay the same if you put the same amount in each time.

After you have finished your 20 turns, you will play one further series of 20 turns but with a new, and potentially different '**black box**'. The two boxes may or may not have the same mathematical function as each other, but the functions will always contain a random component, and the functions will always remain the same for the 20 turns. You will be told when the 20 turns are finished and it is time to play with a new black box.

If you are unsure of the rules please hold up your hand and a demonstrator will help you.

I confirm I understand the instructions (click to confirm)

End of instructions.

Standard and enhanced

Here, we present the instructions for the rate of return $e = 1.6$ (one example of an instruction slide is given in Figure 7). The same instructions apply to standard and enhanced treatments. Equivalent instructions apply for the rate of return $e = 6.4$.

Beginning of instructions.

Instructions

Welcome! You are about to participate in an experimental study of human decision making. Thank you for your participation in our study. Please pay careful attention to the instructions on the following screens. If you wish to return to a previous screen, press the left arrow key. You are now taking part in an economic experiment. If you read the following instructions carefully, you can, depending on your decisions, earn a considerable amount of money. It is therefore very important that you read these instructions with care.

Everybody has received the same instructions. It is prohibited to communicate with the other participants during the experiment. Should you have any questions please ask us by raising your hand. If you violate this rule, we shall have to exclude you from the experiment and from all payments.

During the experiment we shall not speak of £Pounds or Pence but rather of "Credits". During the experiment your entire earnings will be calculated in Credits. At the end of the experiment the total amount of Credits you have earned will be converted to Pence at the following rate: 100 Credits = 15 Pence. In total, each person today will be given 3,200 credits (£4.80) with which to make decisions over two economic experiments and their final totals, which may go up or down, will depend on these decisions.

We are researching the decisions people make.

This part of the experiment is divided into separate rounds. In all, this part of the experiment consists of 20 repeated rounds. In each round the participants are assembled into groups of four. You will therefore be in a group with 3 other participants. The composition of the groups will change at random after each round. **In each round your group will therefore probably consist of different participants.**

In each round the experiment consists of two stages. At the first stage everyone has to individually decide how many credits they would like to contribute to a group project. These decisions have consequences for people's earnings. At the second stage you are informed of the contributions of the three other group members to the project and how many credits you have received from the group project. New groups are then randomly formed and the process repeats itself, with everyone again deciding how many credits they would like to contribute. This process will repeat 20 times.

At the beginning of each round, each participant receives 40 credits. In the following we call this your endowment. Your task is to decide how to use your endowment. You have to decide how many of the 40 credits you want to contribute to a group project and how many of them to keep for yourself. The consequences of your decision are explained in detail in the following slides.

Please note:

- The set-up is anonymous. You will not know with whom you are interacting.
- You will interact with a random set of 3 players in each round.
- Your decisions, and your earnings will remain anonymous to other players, even after the session has ended.

We now provide an animated illustration of a hypothetical scenario to demonstrate how the experiment works. In the following demonstration, we will use these green 'disks' to represent 'credits'. 1 disk equals 1 credit.

Example

Remember! Credits = real money, and the more credits a player has at the end of the experiment, the more money that player will receive.

There are 4 players in each group, For the sake of convenience, we refer to these players as Player A, Player B, Player C, and Player D.

These 4 players then each receive an endowment of 40 credits.

Each of the players can then choose to make a contribution to the group project. They can contribute anything from zero to 40 credits. Non contributed credits are kept in the player's private account. They do this at the same time and anonymously.

Each player is then informed of the decisions of all their group members, although no one will know who the players are and they will randomly change in each round.

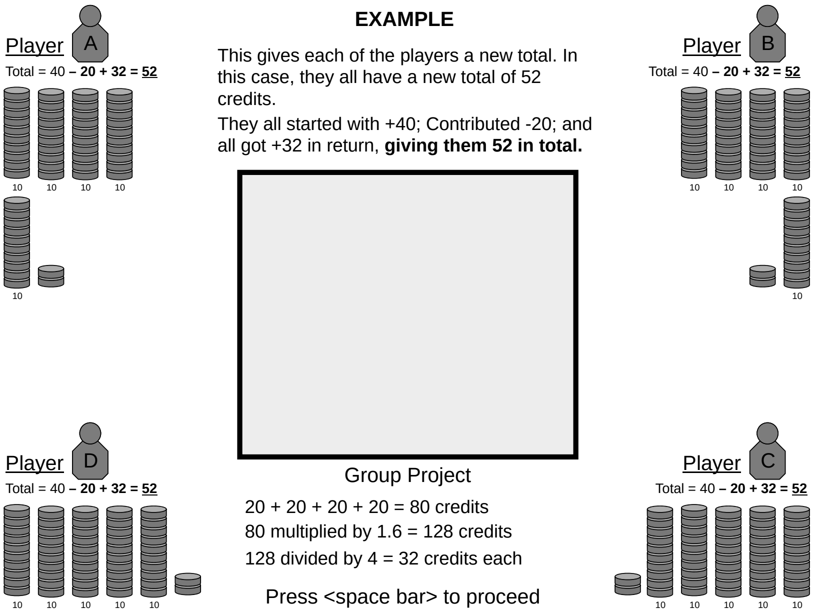
After all 4 players have made their decision to contribute or not, and by how much, the resulting total of contributed credits is automatically MULTIPLIED.

In your experiment, in every round, and in this example here, the total will be multiplied by 1.6. So for an imaginary total of 10 credits, this would result in 16 credits.

This new total is then always shared out equally between all 4 players. So, after the multiplication occurs, each player receives one quarter of the credits that are in the group project.

In this example, each player chooses to contribute 20 of their 40 credits. This means the group project has 80 credits; 20 + 20 + 20 + 20 = 80 credits. Which when multiplied by 1.6 results in 128 credits; 80 multiplied by 1.6 = 128 credits. These 128 credits are then shared out equally, giving 32 credits back to each player; 128 divided by 4 = 32 credits each. This gives each of the players a new total. In this case, they all have a new total

Figure 7: **Standard and enhanced instructions; example slide for** $e = 1.6$**.**



of 52 credits.

They all started with +40; Contributed -20; and all got +32 in return, **giving them 52 in total.**

That is the end of the demonstration.

Remember, this was just one of many possible scenarios. In the rounds you will now play, all players are free to choose how much they wish to contribute to the pot.

End of instructions.

Appendix B: On-screen output in each treatment

The following three tables summarize the post-decision feedback information that participants received in z-Tree under the three treatments: (a) feedback in black-box; (b) feedback screen #1 in standard and enhanced (identical in both treatments); (c) feedback screen #2 in standard and enhanced (dashed lines border the information that was only shown in enhanced).

(a)

GAME SUMMARY	Number of Coins
Initial Coins	40
Minus (-) your input	15
Plus (+) the output returned	24
Your final number of coins	49

This screen lists your decisions and the results, along with your income (for this turn).

(b)

Player Name	Contribution (0-40)
You	15
player A	10
player B	30
player C	5
Total contributions	60
Total after growth	96

This screen lists the decisions of you and the other players (in random order) in your group (for this round). Remember the terms Player A, B, & C **are meaningless** as you play with randomly selected people in each round. The group's total contributions and the new total after the 'growth' stage are also shown.

(c)

Player Name	Contribution (0-40)	Player Income =	Credits retained	+ Credits returned	= Total credits
You	15	For you =	25	24	49
player A	10	player A	30	24	54
player B	30	player B	10	24	34
player C	5	player C	35	24	59

Your income from your group's contributions and subsequent 'growth' is shown.

This screen lists the decisions and earnings of you and the other players (in random order) in your group (for this round).

Remember the terms Player A, B, & C **are meaningless** as you play with randomly selected people in each round.

Appendix C: Regression outputs

Table 3: **Search volatility (black box).**

(a) Test 1: *search volatility* null rejected ($W = 83.6$, $p < 0.01$)	(b) success	(c) failure	(d) total
standard deviation	13.7	11.1	12.6
mean	+0.8	-1.2	-0.2
frequency	2,266	2,198	4,464
Test 2: *rate of return* null not rejected ($W = 0.1$, $p > 0.87$)	$e = 1.6$	$e = 6.4$	total
standard deviation	12.5	12.6	12.6
mean	-0.6	+0.2	-0.2
frequency	2,232	2,232	4,464
Test 3: *phase* null not rejected ($W = 0.3$, $p > 0.59$)	phase 1	phase 2	total
standard deviation	12.7	12.4	12.6
mean	-0.2	-0.1	-0.2
frequency	2232	2232	4464
Test 4: *diminishing volatility* null rejected ($W = 47.0$, $p < 0.01$)	periods 1-9	periods 10-20	total
standard deviation	13.7	11.6	12.6
mean	-0.1	-0.2	-0.2
frequency	1,984	2,480	4,464

We use black box data from phases 1 and 2. We perform Levene's robust variance tests for search volatility (Test 1), differences in distribution for the two rates of return (Test 2), for phases one and two (Test 3), and diminishing volatility (Test 4). W is the Levene's test statistic.

Table 4: **Search breadth (black box)**.

	Coefficient (test statistic)
1 if "failure"	2.73 (9.77)**
1 if "period"<10	2.01 (2.43)*
1 if "phase"=1	2.32 (10.81)**
Group dummies	not significant
Individual fixed effects	not listed
Periods	not significant
Constant	2.42 (3.52)**
Observations	4,464
Adjusted R^2	0.25

*: $p < 0.05$; **: $p < 0.01$.

We use black box data from phases 1 and 2. We perform an OLS regression of absolute adjustments on success-versus-failure adjustments controlling for phase, group, period and individual fixed effects with individual-level clustering.

Table 5: **Inertia (black box)**.

	Coefficient (test statistic)
1 if "failure"	-0.52 (-10.39)**
1 if "period"<10	0.22 (1.37)
1 if "phase"=1	-1.21 (-50.38)**
Group dummies	not significant
Individual fixed effects	not listed
Periods	negative until period 6, not significant thereafter
Cut	0.04 (0.47)**
Observations	4,464
Adjusted R^2	0.40

*: $p < 0.05$; **: $p < 0.01$.

We use black box data from phases 1 and 2. We perform an ordered probit regression of absolute adjustments on success-versus-failure adjustments controlling for phase, group, period and individual fixed effects with individual-level clustering.

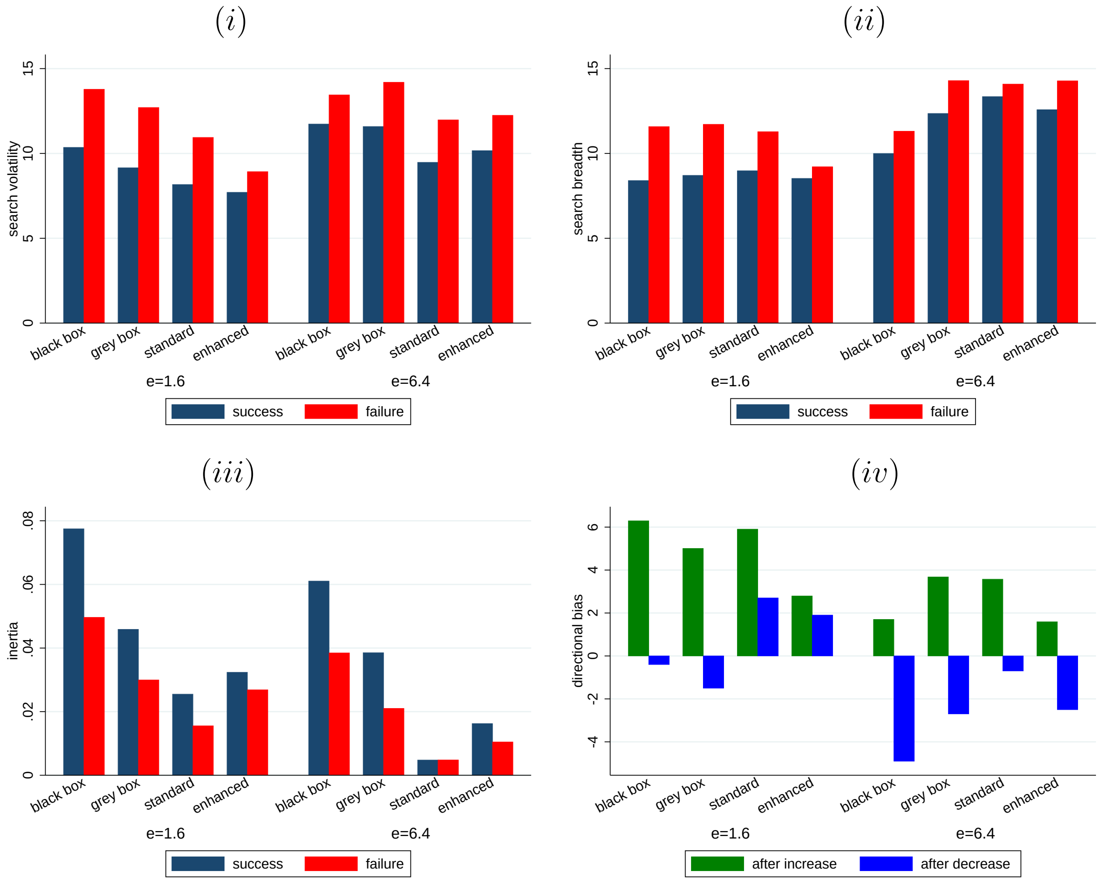
Table 6: **Directional bias (black box)**

	Coefficient (test statistic)
1 if "up"	-10.48 (8.93)**
1 if "up" and "success"	5.06 (5.20)**
1 if "down"	3.62 (2.41)*
1 if "down" and "success"	-3.33 (3.37)**
Individual fixed effects	not listed
Periods	negative, significant
Group dummies	not significant
1 if "phase"=1	-24.90 (13.16)**
Observations	3,222
Adjusted R^2	0.14

*: $p < 0.05$; **: $p < 0.01$.

We use black box data from phases 1 and 2. We perform OLS regressions (without constant) to test for the directional bias of adjustments for each directional success/ failure impulse controlling for phase, group, period and individual fixed effects with individual-level clustering.

Figure 8: SEARCH (all treatments, full details).



The bar charts summarize the four search components in all treatments for both rates of return separately. The respective panels are: (i) search volatility, (ii) search breadth, (iii) inertia, (iv) directional bias.

Table 7: **Search volatility (non-black box).**

Test	success	failure	total
Test 1: *standard* null rejected ($W = 83.1$, $p < 0.01$)			
standard deviation mean frequency	8.9 +1.5 2,314	11.5 -2.4 1,862	10.3 -0.2 4,176
Test 2: *enhanced* null rejected ($W = 30.6$, $p < 0.01$)			
standard deviation mean frequency	9.0 +0.6 2,189	10.7 -1.2 2,131	9.9 -0.3 4,320
Test 3: *grey box* null rejected ($W = 61.3$, $p < 0.01$)			
standard deviation mean frequency	10.5 +1.0 2,045	13.5 -1.2 1,987	12.1 -0.0 4,032

We use non-black box data; phases 3 and 4 for grey box, phases 1-4 for standard and enhanced. We perform Levene's robust variance tests for search volatility in standard (Test 1), enhanced (Test 2), and grey box (Test 3. Recall W is the Levene's test statistic.

Table 8: **Search breadth (non-black box)**.

	(i) **Grey box: Coefficient (test statistic)**	*(ii)* **Standard: Coefficient (test statistic)**	*(iii)* **Enhanced: Coefficient (test statistic)**
1 if "failure"	2.10 (5.10)**	1.64 (4.34)**	1.10 (2.73)**
1 if "period"<10	not significant	not significant	not significant
Phase dummies	"phase"=3, positive	not significant	if "phase"=2, negative
Group dummies	not significant	not significant	not significant
Individual fixed effects	mostly negative	mostly negative	mostly negative
Periods	mostly not significant	mostly not significant	mostly not significant
Constant	16.28(11.44)**	36.83 (24.38)**	12.27 (8.63)**
Observations	2,259	1,861	1,777
Adjusted R^2	0.42	0.41	0.54

*: $p < 0.05$; **: $p < 0.01$.

We use non-black box data; phases 3 and 4 for grey box (panel *(i)*), phases 1-4 for standard (panel *(ii)*) and enhanced (panel *(iii)*). We perform OLS regressions of absolute adjustments on success-versus-failure adjustments controlling for phase, group, period and individual fixed effects with individual-level clustering.

Table 9: **Inertia (non-black box)**.

	Standard & Enhanced coefficient (test statistic)	**Grey box coefficient (test statistic)**
1 if "failure"	-0.23 (2.28)**	-0.29 (2.72)**
1 if "enhanced"	0.02 (0.08)	n/a
Phase dummies	5.11 (26.49)** for "phase"=1 others not significant	5.17 (26.59)** for "phase"=3
Group dummies	not significant	not significant
Individual fixed effects	not listed	not listed
Periods	not significant	not significant
Cut	6.13 (20.12)**	6.48 (20.78)**
Observations	8,464	4,032
Adjusted R^2	0.40	0.34

*: $p < 0.05$; **: $p < 0.01$.

We use non-black box data; phases 1-4 for standard and enhanced (panel *(i)*), and phases 3 and 4 for grey box (panel *(ii)*). We perform an ordered probit regression of absolute adjustments on success-versus-failure adjustments controlling for treatment effects phase, group, period and individual fixed effects with individual-level clustering.

Table 10: **Directional bias (non-black box)**

	Treatment	**Adjustment: coefficient (test statistic)**
1 if "up"	black box	-10.44 (8.83)**
1 if "up" and "success"	black box	5.07 (5.18)**
1 if "down"	black box	3.59 (2.37)*
1 if "down" and "success"	black box	-3.31 (3.34)**
1 if "up"	grey box	-10.61 (7.37)**
1 if "up" and "success"	grey box	6.55 (5.73)**
1 if "down"	grey box	4.69 (3.16)**
1 if "down" and "success"	grey box	-2.73 (2.70)**
1 if "up"	standard	-4.81 (3.86)**
1 if "up" and "success"	standard	9.85 (8.56)**
1 if "down"	standard	10.82 (5.33)**
1 if "down" and "success"	standard	-0.06 (0.04)
1 if "up"	enhanced	-5.96 (3.82)**
1 if "up" and "success"	enhanced	3.25 (2.49)*
1 if "down"	enhanced	8.04 (4.10)**
1 if "down" and "success"	enhanced	-1.10 (0.76)
Group dummies		not significant
Phase dummies		positive, significant
Period dummies		not significant
Treatment dummies		black box negative, others positive
Return rate dummies		negative, significant
Individual fixed effects		not listed
Observations		9,119
Adjusted R^2		0.15

*: $p < 0.05$; **: $p < 0.01$.

We use all the data. We perform OLS regressions (without constant) to test for the directional bias of adjustments for each directional success/ failure impulse in each treatment controlling for phase, group, period and individual fixed effects with individual-level clustering.