

# Collective Experimentation: A Laboratory Study

Mikhail Freer<sup>a</sup>, César Martinelli<sup>b,\*</sup>, Siyu Wang<sup>c</sup>

<sup>a</sup>*ECARES, Université Libre de Bruxelles, Av. Franklin D. Roosevelt 50, CP 114/04, 1050 Brussels, Belgium*

<sup>b</sup>*Interdisciplinary Center for Economic Science, George Mason University, 4400 University Drive, MSN 1B2, Fairfax, VA 22030, USA*

<sup>c</sup>*Department of Economics, Missouri State University, 901 S. National Ave., Springfield, MO 65897, USA*

---

## Abstract

We develop a simple model of collective experimentation and take it to the lab. In equilibrium, as in the recent work of [Strulovici \(2010\)](#), majority rule has a bias toward under-experimentation, as good news for a minority of voters may lead a majority of voters to abandon a policy when each of them thinks it is likely that the policy will be passed by a future majority excluding them. We compare the behavior in the lab of groups under majority rule and under the optimal voting rule, which precludes voting in intermediate stages of the policy experiment. Surprisingly, simple majority performs better than the (theoretically) optimal voting rule. Majority rule seems to lead to better outcomes than other forms of voting when players make mistakes.

*Keywords:* collective experimentation, strategic voting, over-optimism

*JEL:* C92, D72, D83

---

## 1. Introduction

Groups and societies often face decisions whose consequences for different individuals are uncertain and can only be learned over time. As an illustration, trade reforms, changes in immigration or environmental policies, big overhauls of the public health or the tax system, etc. often have consequences that are heterogenous and hard to forecast. Policy innovation under those circumstances is akin to an experiment conducted by society which may result in the new policies being entrenched or abandoned as evidence mounts up. Such policy processes differ from the traditional, static collective choice problem, since majority preferences in favor or against the new policies may change over time. A decision rule for

---

\*Corresponding author

*Email addresses:* [mfreer@ulb.ac.be](mailto:mfreer@ulb.ac.be) (Mikhail Freer), [cmarti33@gmu.edu](mailto:cmarti33@gmu.edu) (César Martinelli), [siyuwang@missouristate.edu](mailto:siyuwang@missouristate.edu) (Siyu Wang)

We thank Alessandra Casella, an anonymous referee, and participants in several audiences for comments and suggestions.

the society must specify not only a voting rule but also how this rule may change over time. In this paper, we develop a simple model of the dynamic collective experimentation problem. We use lab evidence to explore the efficiency implications of different dynamic voting rules from a utilitarian perspective, and compare game theoretic predictions with the behavior of subjects in the lab.

In the model, group members must decide whether to adopt or not a risky policy. It is unclear ex ante who will benefit and who will lose from the policy. Individuals may learn that they will win from the policy in two subsequent stages, after which they have to decide whether to stick to the policy for good or abandon it. Under majority voting at every stage, the policy may, in equilibrium, be derailed in an intermediate stage if a minority learns that they will be winners, as voters in the remainder of the society may anticipate the policy could be adopted in the final stage even if they emerge individually as losers. This is the same bias toward under-experimentation previously identified by [Strulovici's \(2010\)](#) influential contribution. To avoid under-experimentation, it is necessary to increase the threshold for abandoning policy in the intermediate stage. Thus, in the context of our model, optimal voting requires unanimity to derail the risky policy at the intermediate stage, and simple majority in the final stage.

The lab implementation reveals, surprisingly, that adopting simple majority throughout yields better outcomes than the (game theoretic) optimal voting rule. The reason is that voters are sometimes overoptimistic, in the sense that they have a bias in favor of the policy innovation against the available evidence. Under either simple majority or the (game theoretic) optimal voting rule, overoptimism leads to over-implementation in the final stage. At the intermediate stage, overoptimistic behavior reduces inefficient under-experimentation, which is the problem the game-theoretic optimal voting rule was supposed to address. In this context, simple majority is useful because it derails policies that could be adopted later on for good by mistake. We investigate whether the individual bias toward overoptimism may be persistent or temporary, and present evidence consistent with the latter.

The efficiency comparison of unanimity with simple majority is, of course, a classical theme in political economy, at least since the publication of the *Calculus of Consent* ([Buchanan and Tullock, 1962](#)). To the traditional trade-off between majority and minority interests, predicated in a static setting, our work adds a dynamic dimension. Evidence from the lab indicates that repeated majority voting is self-correcting in a way that is not captured by equilibrium analysis. In our setting, majority voting at an intermediate stage partially offsets overoptimism in the last stage. Subjects' voting behavior is consistent with them overestimating their probability of benefiting from the policy after failing to receive evidence supporting that.

There is by now a growing theoretical literature on collective experimentation problems.

[Fernandez and Rodrik \(1991\)](#) provide the seminal contribution, showing that a policy may have minority support ex ante even if it will have majority support ex post for sure if individual voters are afraid of ending up in the losing minority. The above-mentioned article by Strulovici models collective experimentation as a continuous time game with a continuum of voters, identifies the deviations from efficiency implied by adopting majority voting at every moment in time, and derives the (time and state contingent) optimal voting rule. Our model recovers the strategic elements of [Strulovici's \(2010\)](#) model in a discrete time setting with few voters in a way that can be taken to the lab. This allows to check whether the strategic incentives identified by the literature operate in a controlled setting, and explore deviations from equilibrium predictions and their consequences.

Other recent theoretical work on collective experimentation has focused on collective search by committees ([Albrech et al., 2010](#); [Moldovanu and Shi, 2013](#)), on optimal voting rules for two period models ([Messner and Polborn, 2012](#)), on incentives to over-experiment for preemptive reasons when the policymaker may not remain in power ([Callander and Hummel, 2014](#)), on dynamic sequential acquisition by committees ([Chan et al., 2015](#)), and on extending Strulovici's model to consider bad signals ([Khromenkova, 2017](#)). In spite of the growing interest and relevance of this line of work, we are not aware of other lab research on dynamic collective experimentation.

The remainder of this paper organized as follows. Section 2 presents the simple collective experimentation model we use and provides theoretical predictions. Section 3 provides details about the experimental design. Section 4 presents results obtained in the lab. Section 5 gathers final remarks. All omitted proofs are collected in the Appendix.

## 2. Theoretical framework

We consider a committee with three members (players) that make decisions over three periods. Players can be of high or low type; players do not know their own types nor the types of other players. In the first period, players vote on whether to start experimenting with a risky project or not. If the committee starts experimenting, each player receives a public signal. The signal can be *high* or *uncertain*. Each player receives the high signal with probability  $q \in (0, 1)$  if the player is of high type and with probability 0 if the player is of low type. That is, high signals are conclusive evidence of being of high type. If players do not start experimenting, the risky project is abandoned and an alternative safe project is implemented.

In the second period, after observing the public signals received by all players, players vote on whether to continue experimenting with the risky project. If the committee continues, players receive signals again. For notational simplicity, since their type has been already disclosed, we assume that players who received high signals in period 1 receive them again

in period 2. Each of the other players, as in the previous period, receive the high signal with probability  $q \in (0, 1)$  if they are of high type and with probability 0 if they are of low type. We denote by  $w_t \in \{0, 1, 2, 3\}$  the total number of high signals received in period  $t \in \{1, 2\}$ , with  $w_2 \geq w_1 \geq 0$ . If the committee does not continue experimenting, the risky project is abandoned and the safe project is implemented.

Finally, in the third period, if the committee has experimented so far with the risky project, players vote on whether to implement the risky project or the safe project. Afterwards, all payoffs are realized. If a player is of high type and the risky project is adopted, the player receives a payoff of  $h$ . If a player is of low type and the risky project is adopted, the player receives a payoff of  $l$ . If the safe project is adopted instead, all players receive a payoff of  $s$ , where  $l < s < h$ .

Denote by  $p_0$  the initial probability that the player is of a high type. Bayes' Law implies that the probability that a player is of high type after receiving an uncertain signal in period 1 is

$$p_1 = \frac{p_0(1 - q)}{1 - p_0q},$$

and the probability that a player is of high type after receiving uncertain signals in periods 1 and 2 is

$$p_2 = \frac{p_1(1 - q)}{1 - p_1q} = \frac{p_0(1 - q)^2}{1 - 2p_0q + p_0q^2}.$$

Denote by

$$r = p_2h + (1 - p_2)l$$

the expected payoff from the risky project for a player that received uncertain signals in periods 1 and 2. We assume  $r < s$ , so that a player who does not receive high signals would prefer to adopt the safe project.

Further we present the optimal solution of the experimentation game from a simple utilitarian perspective. Next, we describe equilibrium behavior under simple majority voting in every period. In consonance with [Strulovici \(2010\)](#), we find a parameter region under which simple majority voting would result in under-experimentation compared to the socially optimal level after partial good news. Finally, we present a voting procedure which allows to overcome the problem of under-experimentation. Our solution concept is Perfect Bayesian Equilibrium in undominated strategies. As customary in voting games, we eliminate weakly dominated strategies to avoid trivial equilibria in which no player is decisive. For notational simplicity, we assume that players vote to continue experimenting when indifferent, and likewise the social planner breaks indifference in favor of the risky project.

### 2.1. Social planner

A utilitarian social planner would adopt the risky project in the third period if and only if  $w_2 h + (3 - w_2)r \geq 3s$  or equivalently if and only if

$$\frac{h - s}{s - r} \geq \frac{3 - w_2}{w_2}.$$

The condition under which the decision of the social planner coincides with the outcome of the majority voting in the last period is immediate:

**Proposition 1.** *If*

$$\frac{1}{2} \leq \frac{h - s}{s - r} \leq 2,$$

*then the optimal decision from a simple utilitarian perspective is to adopt the risky project if and only if  $w_2 \geq 2$ .*

The proposition is a straightforward application of the condition derived above. Note that the decision of the social planner depends on the total number of high signals received in the second period and not on how many of those were received in the first period. Figure 1 summarizes the decisions made by the social planner. Nodes correspond to the decisions made by the social planner, gray lines to possible actions for the social planner, and thick lines correspond to the optimal actions. The social planner starts experimenting with the project and chooses to continue no matter the number of high signals in the first period. However, the project is implemented if and only if there are at least two high signals received after the second period.

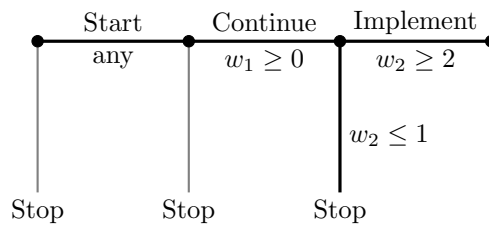


Figure 1: Social planner decision tree

### 2.2. Majority voting

Next we provide conditions under which, in equilibrium, repeated majority (majority voting every period) deviates from the utilitarian social planner's choices by stopping the risky project after a minority gets good news the first period, but otherwise coincides with the social planner choices (see Figure 1). We proceed by backward induction, from the last period to the first. Since by assumption  $h > s > r$ , in the third period only players who have

obtained a high signal vote to implement the policy, so the policy is implemented if  $w_2 \geq 2$ . In the second period players choose the outcome which delivers them the highest expected continuation utility given the updated beliefs after receiving the first set of signals. In the first period players make choices given their expected continuation utility derived from prior beliefs.

**Proposition 2.** *Assume repeated majority. If*

$$\frac{1 - p_1q}{2 - p_1q} \leq \frac{h - s}{s - r} < 1 - p_1q,$$

*then the risky project is adopted if and only if  $w_1 \neq 1$  and  $w_2 \geq 2$ .*

Under the conditions of Proposition 2, the project is implemented if and only if there are at least two high signals received after the second period *and* there is no minority getting good news in the first period. That is, along the lines of [Fernandez and Rodrik \(1991\)](#), the project is stopped in the second period even if it is likely to be supported by a majority in the third and final period (one high signal received the first period). As in [Strulovici \(2010\)](#), continuation of the policy experiment is not monotone in the number of supportive signals obtained in an intermediate stage. That is, the project is continued if there is either no high signal or at least two high signals are received in the second stage, and the project is abandoned if there is exactly one high signal received in the second stage.

Figure 2 illustrates collective choice outcomes under majority voting.

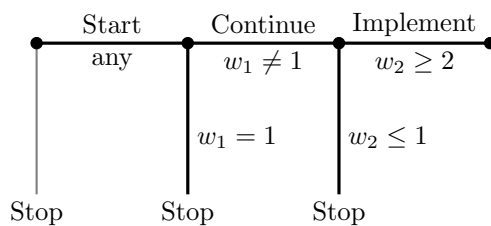


Figure 2: Majority decision tree

### 2.3. Optimal voting rule

Finally, we present a voting rule which delivers the socially optimal outcome in equilibrium. Recall that under-experimentation appears if subjects receive only one high signal after the first period. An intuitive solution to overcome this problem is to eliminate the voting in the second period. Hence, an optimal voting rule is as follows. Players vote to start experimentation with the project and the decision is made according to majority. Next, they receive two sets of signals ( $w_1$  and  $w_2$ ) without making any decision until the third period.

Finally, players vote to implement or stop the project and the decision is made according to majority.

**Proposition 3.** *Assume the optimal voting rule. If*

$$\frac{1}{2} \leq \frac{h-s}{s-r} < 1 - p_1q,$$

*then the outcome coincides with the outcome delivered by the utilitarian social planner, and the risky project is adopted if and only if  $w_2 \geq 2$ .*

Figure 3 presents the collective decision tree for the optimal voting rule. In the second period, experimentation is continued automatically if it is initiated, regardless of  $w_1$ .

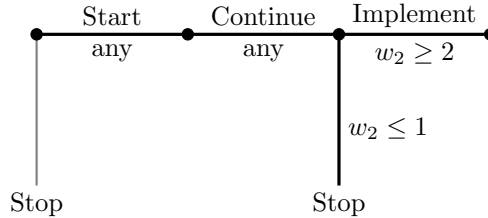


Figure 3: Optimal rule decision tree

Note that the conditions of Proposition 3 are the intersection of those of Propositions 1 and 2. Under these conditions, repeated majority would result in inefficient under-experimentation, while the optimal voting rule would implement the utilitarian solution. Finally, note that the tie-breaking rule players use to make decisions in case of indifference may seem important. However, unless the parameter conditions in the propositions above are satisfied as equalities, the tie-breaking rule has no more than notational significance.

### 3. Experimental design

We conduct a lab experiment based on the theoretical model from Section 2 with the parameter values described in Table 1. Under the parameters of Table 1, the updated probability of being of the high type after uncertain signals is  $p_1 = 1/3$  in period 1 and  $p_2 = 1/5$  in period 2, and the expected value of the risky project after two uncertain signals is  $r = 140$ . The ratio  $(h-s)/(s-r)$  of gains over losses of implementation takes the value  $5/7$ , satisfying the conditions stated in Proposition 3. Equilibrium predicts that simple majority voting in each period would be suboptimal, while the optimal voting rule delivers outcomes that coincide with the utilitarian planner choices. Note that the theoretical predictions are based on the assumption that players are risk-neutral. If players were risk-averse, they would have a larger incentive to under experiment under majority, furthering the advantage

of the optimal mechanism. (If players were extremely risk-averse, they would not initiate any experimentation, eliminating any advantage of the optimal mechanism.)

Parameter	Value
Probability of being high type $p_0$	1/2
Probability that signal reveals the type $q$	1/2
High payoff $h$	500
Low payoff $l$	50
Safe payoff $s$	350
Exchange rate	20 tokens per US dollar

Table 1: Parameters specification

The experimental design consists of two treatments. In the first treatment subjects make decisions under simple majority in every period, and in the second treatment subjects make decisions under the optimal voting rule described above. We refer to the treatments as *repeated majority* and *optimal rule*. The game was played 15 rounds with random reshuffling of the groups, to ensure that there was no repetition of the same group composition. The experiment was coded and conducted using oTree (Chen et al., 2016). Experimental instructions and screenshots of the interface can be found in an online appendix. We conducted the experiment with 96 undergraduate students from George Mason University. There were 48 subjects in the repeated majority treatment and 48 in the optimal rule treatment.<sup>2</sup> Subjects' earnings varied between \$10 and \$40.

#### 4. Experimental results

We start by looking at group level results to test the theoretical predictions and compare the performance of the two voting rules. Then we move on to analyzing individual behavior to explore possible biases which caused the observed group level behavior. Next, we explore possible learning patterns. Finally, we present results from a structural estimation using a version of the Quantal Response Equilibrium model (introduced by McKelvey and Palfrey, 1995, 1998).

As a convention, in what follows we refer to *over-* or *under-experimentation* when talking about the decisions in the second period, and we refer to *over-* or *under-implementation* when talking about the decisions in the third and last period. Note that whenever we talk about

---

<sup>2</sup> We conducted four sessions for each treatment, with twelve subjects in each session. Five sessions (three for the repeated majority and two for the optimal rule treatment) were conducted in April 2017, and three additional sessions (one for repeated majority and two for the optimal rule treatment) were conducted in May 2018.



over-/under-experimentation or over-/under-implementation we are making a comparison with the socially efficient levels.

#### 4.1. Group level analysis

Figure 4 summarizes the results. A first observation is that subjects (in both treatments) start the project with very high probability, in line with theoretical predictions.

Recall that in equilibrium majority voting deviates from the optimal utilitarian solution in the second period if one high signal is received. In this case, equilibrium predicts that the policy will be abandoned, while optimality requires to continue experimenting. In the lab, the risky policy is abandoned with a frequency of 33%, which is significantly lower than the equilibrium prediction. At the same time, the observed frequency of stopping is significantly higher than in the optimal solution. There is also significant under-experimentation in the case of no high signals received: if two or more high signals were received, the policy is terminated with a frequency of 11%.

We do not find significant differences in stopping probabilities in the second period after receiving one or zero high signals (similar evidence can be found in Table 2). In fact, we obtain continuation probabilities that are monotonic in the number of high signals received, which goes against the predictions of the theoretical model. In particular, it contradicts the prediction of Proposition 2.

Numbers in square brackets at the final leaves of the game tree in Figure 4 show the frequency of policy implementation for the different possible trajectories of high signals received. We can see that the optimal voting rule implements significantly more frequently the risky policy when it is efficient to do so ( $w_2 \geq 2$ ). However, the optimal voting rule also implements significantly ( $p < .05$ ) more frequently the risky policy when it is *inefficient* to do so ( $w_2 < 2$ ). Neither of the two treatment effects is eliminated with learning.<sup>3</sup>

	$w_2 = 0$	$w_2 = 1$	$w_2 \geq 2$
Majority treatment (if continued)	12.90%	36.67%	90.67%
Majority treatment (in total)	7.84%	22.45%	74.73%
Optimal treatment	11.76%	36.47%	94.25%

Table 2: Implementation probabilities by  $w_2$

In Table 2 we aggregate implementation frequencies by  $w_2$  only. The first line corresponds to the implementation probability for majority treatment if the group decided to continue at previous stages and therefore observed the public signals. Note that implementation

<sup>3</sup> In order to control for the possibility of learning we have conducted a robustness check, splitting the data from all sessions in three parts: the first, middle, and last five rounds. We do not find different qualitative patterns after splitting the data.

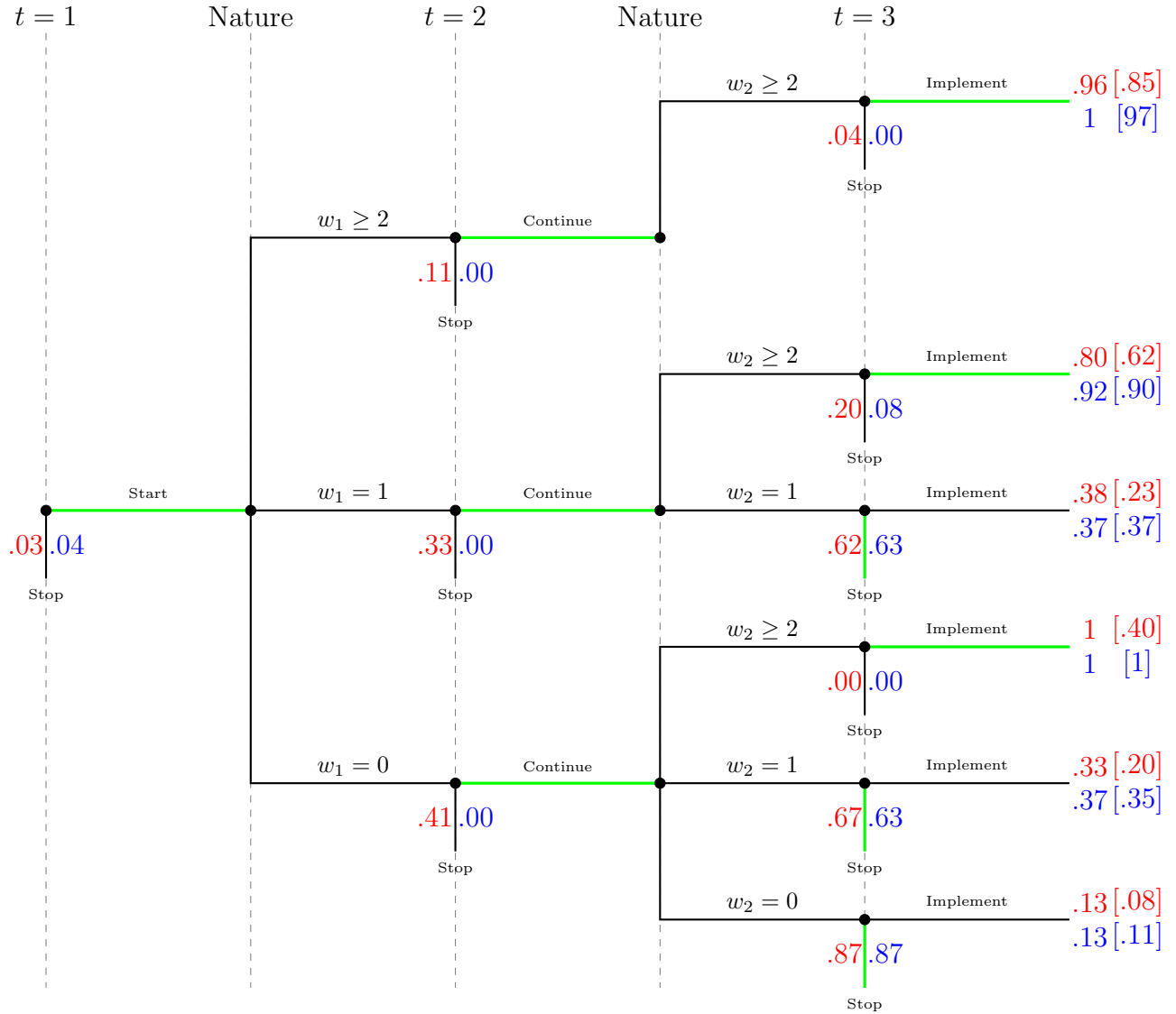


Figure 4: Probabilities of abandoning and implementing the policy on the collective decision tree. Red (on the left of vertical lines and above the horizontal ones) and Blue (on the right of vertical lines and below the horizontal ones) numbers correspond, respectively, to the frequencies obtained in the repeated majority treatment and in the optimal rule treatment. Green lines correspond to the socially optimal decisions at every node. Stopping frequencies and frequencies of policy implementations displayed are conditional on the probability of reaching the node. In square brackets we show probabilities of implementing the policy conditional on the history of signals received.

probabilities in this case are not different from those in optimal treatment in the third line. The second line corresponds to the total probability for majority treatment, i.e. it includes projects which were abandoned at previous stages. If we include all projects, there is significantly less implementation under majority for all possible signals, including  $w_2 = 2$ .

We do not find evidence that significant over-implementation is caused by different collective actions in different treatments at the third period.<sup>4</sup> There are no significant differences at either level of aggregation, as presented in Figure 4 and Table 2. Numbers below the leaves on Figure 4, corresponding to stopping the policy in the third period, show the frequency of abandoning the policy in the last period conditioning on continuing the project at the second period. There are no significant differences over treatments in these frequencies. Similarly we do not find any significant differences at the higher level of aggregation (only by  $w_2$ ) presented on Table 2, when comparing first and third rows. Hence, we cannot reject the hypothesis that group actions are the same in the last period.<sup>5</sup> Hence, over-implementation of policy in the optimal treatment is caused by the forced continuation of the policy.

We compare the expected revenues from each voting rule given the observed patterns at the group level. Note that it is not sufficient to compare the average welfare in each treatment, because of different realizations of types and signals which may lead to small sample bias. In order to overcome this problem, we conduct a two-step bootstrap procedure with re-weighting the realized outcomes according to the theoretical distribution of outcomes.<sup>6</sup>

Details of the procedure as well as detailed evidence on empirical and theoretical distribution of types and signals can be found in the Appendix B. We report the normalized group welfare–total earnings of the group from the implemented option minus the total earnings of the group if the safe option had been implemented. Mean expected earnings are  $-3.03$  (median  $-2.74$ ) in the repeated majority treatment and  $-14.43$  (median  $-14.25$ ) in the optimal rule treatment. Differences are significant with  $p < 0.001$ . We conclude that repeated majority voting delivers higher expected welfare than the optimal voting rule.

#### 4.2. Individual level analysis

Next, we perform the analysis of the individual behavior in order to investigate the drivers of the observed behavior at the group level. Recall that the equilibrium action in

---

<sup>4</sup>Note that there may be the selection at the individual level behavior between treatments although, as shown, it does not lead to different group outcomes. We investigate individual behavior and selection in the next subsection while examining individual-level strategies.

<sup>5</sup> Although there may be an issue regarding the power of our tests, we have 240 observations for each treatment. Moreover, for  $(w_1, w_2) = (1, 1)$ , the stopping frequency is higher in the optimal rule treatment, which can serve as additional evidence that there is no difference in strategies used at the last period.

<sup>6</sup> It may seem that such a correction is unnecessary given that we have 720 observations for each treatment. However, the realization of signals is a random variable that can take one out of ten values. Even a small measurement error may add up to significantly bias the total welfare we obtain (see Appendix B for details).

both treatments in the first period is to vote to start the risky project, and in the third period it is to vote to implement the risky project for good if the player has obtained a high signal, and against otherwise. In fact, in the third period it is a weakly dominant strategy to vote with the signal. In the second period, in the repeated majority treatment, under our parameter specification, the equilibrium action is to vote to continue the project if a player has obtained a high signal or if no one else has obtained a high signal, and to vote to abandon the project otherwise.

	Majority						Optimal		
	$t = 2$			$t = 3$			$t = 3$		
	$s_i = 0$	$s_i = 1$	any	$s_i = 0$	$s_i = 1$	any	$s_i = 0$	$s_i = 1$	any
$w_t = 0$	.539 (.472 – .603)	–	.539 (.472 – .603)	.785 (.688 – .825)	–	.785 (.688 – .863)	.773 (.706 – .830)	–	.773 (.706 – .830)
$w_t = 1$	.448 (.379 – .517)	.733 (.638 – .815)	.543 (.486 – .599)	.750 (.663 – .825)	.833 (.715 – .917)	.778 (.710 – .836)	.741 (.667 – .806)	.855 (.761 – .923)	.779 (.722 – .829)
$w_t \geq 2$	.520 (.374 – .663)	.866 (.789 – .923)	.759 (.689 – .823)	.702 (.566 – .816)	.911 (.856 – .949)	.858 (.805 – .901)	.587 (.467 – .699)	.956 (.914 – .981)	.847 (.797 – .889)
any	.499 (.454 – .544)	.802 (.743 – .853)	.595 (.558 – .632)	.752 (.696 – .802)	.890 (.842 – .928)	.815 (.778 – .848)	.728 (.683 – .769)	.924 (.885 – .953)	.802 (.771 – .831)

Table 3: Frequency of equilibrium actions with 95% confidence intervals in the parenthesis

Table 3 shows the probability of playing equilibrium actions. Every block of the table presents a given treatment at a given period of the game. Every cell reports the corresponding probability given the signal received by the subject and given the number of high signals received by the group. Probability of equilibrium behavior in the third period on average is about 80% in both treatments. Moreover, there is no significant difference between treatments in any of the cases and levels of aggregation. This result can serve as evidence to confirm that there is no difference in subjects' strategies in the third period.

Comparing the probability of equilibrium play between periods within majority treatment, we find that it is lower in the second period (59.5% vs. 81.5%). This difference is not significant if the subject received a high signal (80.2% vs. 89%), but it is very clear when the subject received an uncertain signal (49.9% vs. 75.2%). Conditioning this difference on the number of high signals received by the group, it is significant if no more than one high signal was received.<sup>7</sup> Hence, we see that the quality of decision making is lower in the second period and depends on the signal the subject received.

We observe subjects largely deviating from theoretical predictions in the case of the group receiving only one high signal. This deviation primarily comes from the fact that if a subject receives an uncertain signal, she still votes for the policy with high probability. This particular behavior helps overcome the theoretically predicted inefficiency. Moreover, we

<sup>7</sup> The difference is not significant if two or more high signals were received, but it is likely the effect of the small size of the subsample.

observe that in general subjects deviate from equilibrium behavior significantly more after an uncertain signal than after a high signal. We refer to this phenomenon as overoptimism—subjects tend to vote for the risky project regardless of the signal received.

One can ask whether there is an alternative explanation for the observed overoptimistic behavior. The usual suspects, in this case, are risk aversion and social preferences. However, neither of them can explain the overoptimistic behavior and be consistent with the data in other cases as well. We argue first that risk aversion should lead to under-experimentation. To see this, recall that abandoning the policy delivers a riskless option, while continuing with the policy is riskless if and only if subject received a high signal. Consider the choice in the third period. If the subject received a high signal, then she should vote for the risky project no matter her risk preferences, since the payoff is deterministic in both cases and the payoff from the risky project is simply greater. If the subject only received uncertain signals, then she faces a choice between the risky project which delivers a non-degenerate lottery and a safe project which is riskless. In this case, even risk neutral subjects would prefer the safe option. Hence, risk-averse subjects would do the same. Given that the collective action in the third period is identical for risk neutral and risk averse subjects, they face the same lotteries in the second stage. Hence, if the risk-neutral subject chooses to vote against the risky project upon receiving an uncertain signal (see Lemma 2), the risk-averse subject would do the same. Therefore, the risky project should be abandoned if  $w_1 \leq 1$ , which is the equilibrium prediction for risk-neutral subjects.

Now consider inequality averse preferences (see e.g. [Fehr and Schmidt, 1999](#); [Bolton and Ockenfels, 2000](#)). Assume that subjects are risk neutral and care about their own earning and the inequality between players. Note that inequality is equal to zero if all players receive equal payoffs, and strictly positive if payoffs are unequal. Thus, for any subject satisfying the premises of the generalized inequality aversion model (see [Dolgoplov and Freer, 2018](#)), there will be a bias toward the riskless alternative. Decisions in the third period would depend on the particular utility function, and the degree of inequality aversion of subjects. Consider a subject who receives an uncertain signal in the second period. If the project is abandoned, there is no inequality between players, while if the project is continued there is strictly positive expected inequality unless the risky project is implemented only if  $w_2 = 3$ . Moreover, the subject's expected payoff from the continuation is also lower than the payoff from abandoning the project unless the risky project is implemented only if  $w_2 = 3$  (see Lemma 2). Hence, unless the project is implemented only if every subject received a high signal in the third period, an inequality averse subject would prefer to stop the project in the second period if she received an uncertain signal. Figure 4 shows that there is a strictly positive probability of implementing projects as long as  $w_2 \geq 2$ . Therefore, if the subject has precise enough beliefs about voting outcomes in the third period, inequality aversion would

also result in under-experimentation.

Finally, let us remark about subjects with altruistic preferences. Altruistic enough subject should exhibit the behavior which is close to the behavior of utilitarian social planner. Hence, we should observe frequent voting against the implementation of a risky policy if  $w_2 = 1$ , since implementing the risky policy is suboptimal in this case. However, such behavior pattern is not consistent with observed behavior.

### 4.3. Learning

The observed effects can potentially disappear over time because of learning. We explore this possibility using a comparison between equilibrium play frequencies in the first and last five periods. For robustness, we present the results of a probit regression of the linear learning trends in Appendix C.

	Majority						Optimal		
	$t = 2$			$t = 3$			$t = 3$		
	$s_i = 0$	$s_i = 1$	any	$s_i = 0$	$s_i = 1$	any	$s_i = 0$	$s_i = 1$	any
$w_t = 0$	.181(↓)	–	.181(↓)	.833(↑)	–	.833(↑)	.045(↑)	–	.045(↑)
$w_t = 1$	.861(↓)	.677(↓)	.850(↓)	.001(↑)	.571(↑)	.002(↑)	.043(↑)	.201(↓)	.341(↑)
$w_t \geq 2$	.593(↑)	.513(↑)	.896(↑)	.751(↓)	.274(↑)	.751(↑)	.812(↑)	.497(↑)	.838(↑)
any	.618(↓)	.290(↓)	.506(↓)	.137(↑)	.172(↑)	.025(↑)	.012(↑)	.521(↓)	.073(↑)

Table 4: Results of analysis of effect of learning over the frequency of equilibrium action. We report  $p$ -values are reported in every cell of the table comparing first and last five periods. Arrows in parenthesis report the direction of change from first to last periods of the experiment:  $\uparrow$  represents an increase of the equilibrium frequency;  $\downarrow$  represents a decrease.

Table 4 reports the results of the analysis of learning behavior. Every cell reports the  $p$ -value for the corresponding case. In parenthesis we report the direction of change of the frequency of best-responding,  $\uparrow$  means that frequency of best-responding increased from first to last five periods and by  $\downarrow$  if this frequency decreased. Note that the only significant trends we observe correspond to the actual learning, i.e. increasing probability of equilibrium play.

We do not observe any learning in the second period under the repeated majority treatment. This implies that over-experimentation does not disappear with learning. In particular, we do not observe any learning in the case of  $w_1 = 1$  and  $s_i = 0$ , that is the case in which subjects are deviating significantly from equilibrium behavior. Similarly, we do not observe learning in the case of no high signals received. Thus, the monotonicity of the probability of continuation on the number of high signals persists over time in the lab, in contrast with equilibrium predictions.

We do observe significant learning in the third period for both treatments. In particular, subjects learn to vote against the risky policy if they received uncertain signals and their vote is likely pivotal because no more than one high signal is received by the group. If

the group did not receive high signals, then there is no significant learning in the repeated majority treatment while there is in the optimal rule treatment. The smaller number of observations for  $w_2 = 0$  in the repeated majority treatment due to the high probability of abandoning the risky policy in the second period (41%) implies that the coefficient is not significant anymore due to the lack of power (93 observations in majority vs 189 in optimal treatment). This learning pattern is not surprising overall and it is consistent with theories of learning, e.g. reinforcement learning (see for instance [Erev and Roth, 1998](#)). It may imply that the inefficiency implied by over-implementation under the optimal rule can be resolved by learning.

	$1 \leq \tau \leq 5$	$6 \leq \tau \leq 10$	$11 \leq \tau \leq 15$
Majority	-19.8001	-5.3097	7.9359
Optimal	-27.3631	-40.0585	-0.8032
<i>p</i> -value	$p < .001$	$p < .001$	$p < .001$

Table 5: Normalized welfare evolution over time

Table 5 presents normalized expected welfare values over time. We estimate the expected welfare using the two-step bootstrapping procedure with probability re-weighting described in Appendix B. We normalize the welfare by the baseline of choosing the safe option (i.e. the payoff of 350). Expected welfare remains greater in the repeated majority treatment than in the the optimal rule treatment, in spite of learning. Note that in the first rounds expected welfare is negative in both treatments due to over-implementation. In the last five rounds we observe that expected welfare is positive in the repeated majority treatment, while it is still slightly negative in the optimal rule treatment.

Using the quantal response model with uniform errors (see next section), we can understand why there is little learning to play the equilibrium action in the second period under repeated majority. Since learning to play the equilibrium action in the third period is too slow, there is persistent over-implementation. Given over-implementation, in the QRE model it is in fact a best-response to vote against the risky policy after receiving an uncertain signal in the second period, regardless of the signals of other players. Thus, learning may lead away from equilibrium in the second period under majority, and voting for the risky policy may be attributed to overoptimism rather than to best-responding. The quantal response model explains a persistent difference in welfare in favor of majority rule as long as we allow for the decision making error to be around 5-10% and being constant over time.

#### 4.4. Structural model

We estimate a simple Quantal Response Equilibrium model with behavioral types, similar to [Elbittar et al. \(forthcoming\)](#). The model has two parameters;  $Q$  is a quantal response

parameter while  $p$  is a type parameter. In each period, with probability  $1 - Q$ , each player uniformly randomly chooses one of two actions (voting for or against the risky policy). With probability  $p$ , a player who is not choosing randomly votes for the project no matter what. We interpret  $p$  as the probability of being (temporarily) overoptimistic.<sup>8</sup>

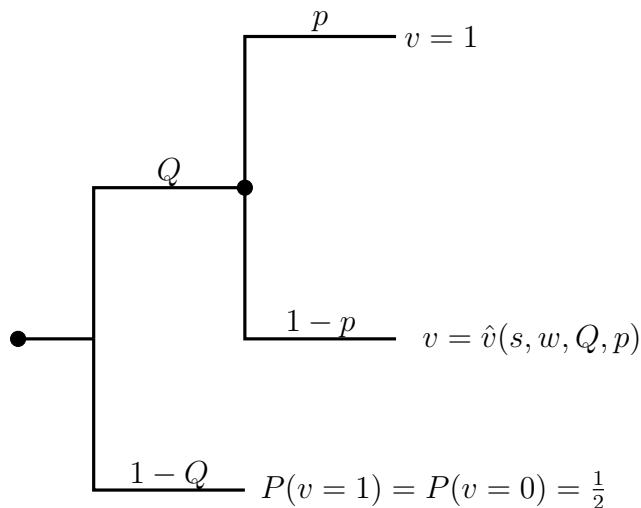


Figure 5: Structural model

Figure 5 illustrates decision-making in the structural model for each period. The term  $\hat{v}_t(s, w, p, Q)$  represents the undominated best-response (voting in favor or against the risky project) for a player in period  $t$  given that the player has received so far a high signal ( $s = 1$ ) or not ( $s = 0$ ), given the high signals received by the group so far  $w \in \{s, \dots, 3\}$ , given the parameters  $p$  and  $Q$ , and given the strategies of other players. The quantal response model we estimate deviates from the original quantal response model of [McKelvey and Palfrey \(1998\)](#) in two main aspects. First, we assume that the distribution of errors does not depend on the payoff delivered by the option. Recall that in the usual (e.g. logit) QRE model the distribution of errors depends on the the average payoff the player would receive from choosing the action. Formally, we make a different assumption about the distribution of errors. There are a few reasons for this. Since we have four cases in the second period and nine cases in the third period, it is necessary to divide the sample into four and nine subsamples correspondingly. Thus, the precision of the average utility estimate is smaller than in a static game, and the estimate of the parameters of the model is noisier. Moreover, if we use the exact expected payoff to estimate the model, we face non-convexities which can

---

<sup>8</sup> We abstain from estimating another popular model of cognitive hierarchies ( $k$ -level) for two reasons. First, incorporating behavioral types into that model is less straightforward. Second, the  $k$ -level model assumes that there are 0-level thinkers, who persistently take random decisions. We do not observe such pattern in the data.



make the estimation computationally intractable. Finally, incorporating behavioral types in the standard quantal response model would be rather problematic. Another difference with the usual QRE model is that instead of  $\lambda > 0$  we use the exact parameter  $Q \in [0, 1]$ , which is a probability of best-responding behavior. This makes the model easier to interpret.

The probability that player  $i$  in period  $t$  votes for the policy ( $v = 1$ ) is

$$P_t(v = 1|s, w, p, Q) = Q((1 - p)\hat{v}_t(s, w, p, Q) + p) + \frac{1 - Q}{2},$$

where  $\hat{v}_t^i(s_t^i, w_t, p, Q)$  is as described above. Best-response depends on the model parameters, since QRE assumes that each player not only deviates from best-response behavior with some probability, but also takes into account that other subjects do so while deciding what is the best-response. Hence, the best-response of the subject is formed under beliefs which would be rational under the null hypothesis (subjects behave according to the quantal response model with overoptimism). That is, players behave as if they know the entire distribution over actions taken by all players. Due to the distribution of errors, as the value of the action increases, we observe a discontinuous increase of the probability of playing the action depending on whether the action is best-response or not. Note that this is unlike the continuous increase of the probability of playing an action in the standard QRE model. The probability that player  $i$  in period  $t$  votes against the policy ( $v = 0$ ) is, in turn,

$$P_t(v = 0|s, w, p, Q) = Q(1 - p)(1 - \hat{v}_t(s, w, p, Q)) + \frac{1 - Q}{2}.$$

Denote by  $n_t^v$  for  $v \in \{0, 1\}$  and  $t \in \{1, 2, 3\}$  the number of subjects who took decision  $v$  in period  $t$ , and let  $\hat{v}_t^i = \hat{v}_t(s_t^i, w_t, p, Q)$  be the estimated action for player  $i$  in period  $t$  conditional on her realized signal  $s_t^i$  and number of signals in her group  $w_t$ . We can derive the following likelihood function for the majority treatment:

$$L(Q, p, n) = \prod_{t \in \{1, 2, 3\}} \left\{ \left( Q((1 - p)\hat{v}_t^i + p) + \frac{1 - Q}{2} \right)^{n_t^1} \left( Q(1 - p)(1 - \hat{v}_t^i) + \frac{1 - Q}{2} \right)^{n_t^0} \right\}.$$

	$Q$	$p$
Majority treatment	.79	.36
Optimal treatment	.75	.05
Pooled data	.77	.07

Table 6: Estimated parameters

Table 6 presents the estimation results for the quantal response model using data from the repeated majority treatment, the optimal rule treatment, and the pooled data. We observe

quite similar estimates for the quality of decision making ( $Q \in [.75, .77]$ ) in the three cases. The estimated overoptimism parameter, however, is much higher in the repeated majority treatment ( $p = .36$  for majority vs  $p = .05$  for optimal). This is explained by the absence of learning not to be overoptimistic in voting in the second period in the repeated majority treatment. Estimates for all cases imply that in the repeated majority treatment the best-response in the second period is the same as in the third period: voting your signal. Splitting the sample in three five-round blocks, we do not observe significant shifts in the estimated coefficient for any of data sets:  $(p, Q) \in [.35, .36] \times [.76, .81]$  for majority treatment data,  $(p, Q) \in [.04, .06] \times [.73, .77]$  for optimal treatment data and  $(p, Q) \in [.06, .07] \times [.77, .79]$  for pooled data.<sup>9</sup> To explore whether individuals behave differently in period  $t = 3$  in the repeated majority treatment than in the optimal rule treatment, we estimate the structural model in the repeated majority treatment treatment dropping behavior in  $t = 2$ . We obtain the estimates  $(p, Q) = (.07, .78)$ , which are very close those for the optimal rule treatment in Table 6.

These estimates are not significantly different from the estimates from the Optimal treatment (that is consistent with result from Table 3). Moreover, note that difference comes in terms of over-optimism, hence, we can treat the further differences as those driven by overoptimistic behavior.

	$w_t = 0$	$w_t = 1$	$w_t = 2$	$w_t = 3$
Majority treatment ( $Q = .79, p = .36$ )	.67	.42	.13	.03
Optimal treatment ( $Q = .75, p = .05$ )	.93	.74	.20	.04
Pooled data ( $Q = .77, p = .07$ )	.93	.73	.18	.03

Table 7: Probabilities of stopping the policy conditional on  $w_t$  high signals

Table 7 presents the stopping probability given the number of signals received by the group. Recall that given our estimates the best-responses in the second and third period are the same—a player should vote for the risky project if and only if the player receives a high signal. We condition stopping probabilities only on the number of high signals received by group. We observe significantly higher probabilities given the estimates for the optimal rule treatment, because it takes into account only voting in the third period which is less subject to overoptimistic behavior.

We can use the structural model to compute counterfactuals. In particular, given that

<sup>9</sup>We observe some variation in coefficients which can be attributed to learning. We observe more variation in the probability of best-responding ( $Q$ ) than in the probability of being overoptimistic ( $p$ ). We do not find any learning regarding overoptimism in the second period, from which a big deal of the estimation of  $p$  comes from. We do not find *significant* shifts, because learning affects mostly  $Q$  and the absolute effect is not that large.

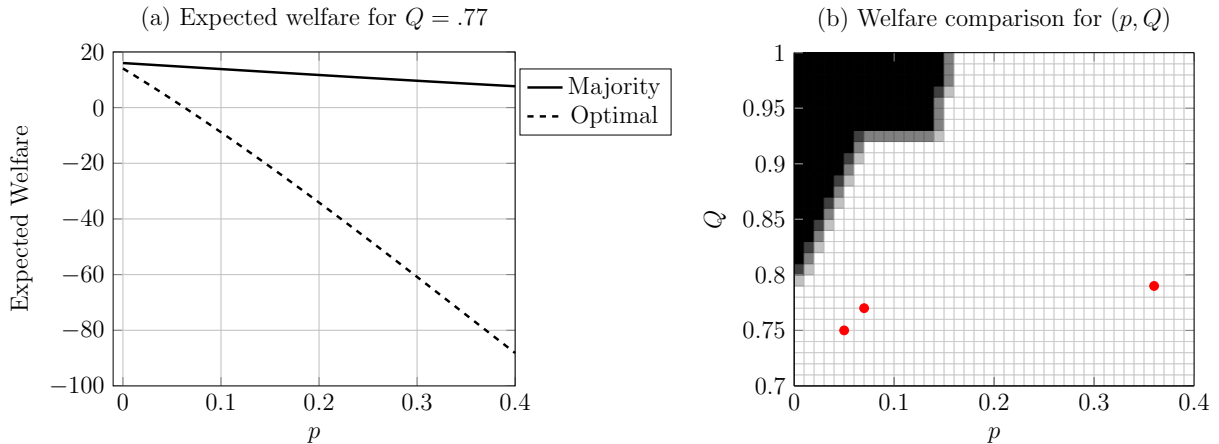


Figure 6: Expected welfare for different levels  $p$  and  $Q$ .

we obtain similar estimates of  $Q$  across treatments, it is illustrative to consider the effects of changes in the level of overoptimism. Figure 6(a) presents the expected welfare estimates for the majority and optimal mechanisms using the quantal response estimate from the pooled data ( $Q = .77$ ) and different levels of overoptimism ( $q$ ). Figure 6(a) shows that expected welfare from the optimal mechanism is lower for every level of overoptimism. This result is consistent with the welfare estimates we presented above.<sup>10</sup>

Figure 6(b) shows which mechanism delivers higher expected welfare for possible levels of  $(p, Q)$ . The shaded region represents the set of parameters for which the optimal mechanism delivers higher expected welfare. Red dots correspond to the positioning of the estimated parameters from Table 6. One can see that either vector of estimated parameters is far from the region in which optimal mechanism delivers higher expected welfare. Moreover, the figure explicitly shows how vulnerable is the optimal mechanism to the optimization and behavioral biases. Even if there is no overoptimism whatsoever, the optimal mechanism delivers higher expected welfare only if the probability of best-responding is at least 80%. Moreover, if the probability of overoptimistic behavior is at least 15%, then simple majority delivers higher expected welfare for any probability of best-responding.

Finally, we attempt to classify subjects in persistent types (best-responder or overoptimistic) using likelihood ratios. For the repeated majority treatment, every subject participated in 15 rounds of the game and made at least once decision, so that for every subject we have between 15 and 45 observations. Similarly, for the optimal treatment, we have between 15 and 30 decisions for every subject.

Persistent best-responder and overoptimistic types have different likelihood functions, and

<sup>10</sup> Moreover, the higher is the level of over-optimism the higher is the difference between the majority and the optimal mechanisms. Similar results hold for other values of  $Q$ ; we present more details in the Appendix.

we compare the odds of them. For all estimated decision making errors the best response is to start a project and to vote further according to the signal player received. Hence, if we denote by  $k_t^v$  the number of times that a given player voted  $v \in \{0, 1\}$  in period  $t$ , we can write down the following likelihood ratio for the player:

$$\lambda(Q, k) = \frac{(Q + \frac{1-Q}{2})^{k_1^1} (\frac{1-Q}{2})^{k_1^0} \prod_{t \in \{2,3\}} \left\{ (Q s_t^i + \frac{1-Q}{2})^{k_t^1} (Q(1 - s_t^i) + \frac{1-Q}{2})^{k_t^0} \right\}}{\prod_{t \in \{1,2,3\}} \left\{ (Q + \frac{1-Q}{2})^{k_t^1} (\frac{1-Q}{2})^{k_t^0} \right\}}.$$

The numerator corresponds to the likelihood that the player is a best-responder given  $Q$ , while the denominator corresponds to the likelihood that the player is overoptimistic given  $Q$ . If  $\lambda(Q, k) > 20$  we classify a player as best responder, if  $\lambda(Q, k) < \frac{1}{20}$  we classify player as overoptimistic, and otherwise the player is unclassified. We interpret 20 : 1 odds, for descriptive purposes, as indicating that with 95% confidence the subject is correctly classified by the estimated model.

	Majority	Optimal	Pooled
best responders	.56	.70	.64
unclassified	.08	.15	.11
overoptimistic	.35	.15	.25

Table 8: Classification of subjects

Table 8 presents the fraction of subject classified by treatment and with pooled data. For every column we use the parameter  $Q$  estimated from the sample (see Table 6). More than half of subjects are classified as best-responders given the estimated parameters. Moreover, even in the optimal treatment (voting only at  $t = 3$  which is less prone for overoptimistic behavior) we see that 15% of subjects can be classified as overoptimistic. Finally, there is non-trivial fraction of subjects who cannot be classified in either of groups (8-15%).

## 5. Final remarks

We construct a simple three-period political economy model of experimentation. In the model, after a group or society starts experimenting with a new policy, individual members may obtain or not conclusive signals that they will be winners from the policy if the policy is adopted permanently. We show that in equilibrium repeated majority rule may have a bias toward under-experimentation, capturing a strategic consequence of collective experimentation first identified by [Strulovici \(2010\)](#). We also describe an optimal voting rule.

The optimal voting rule either requires unanimity for abandoning the policy in intermediate policies, or simply eschews voting in the intermediate period.

We implement the model in the lab with three-player groups. Against equilibrium predictions, we find over-experimentation and, puzzlingly, over-implementation in the lab when there is little evidence in support of adopting the policy under both majority voting and the (theoretically) optimal voting rule. This causes repeated majority voting to outperform the optimal voting rule: allowing voters to vote at intermediate periods about the policy allows them to reduce the probability that the risky policy is adopted for good by mistake in the final period.

We interpret mistakes in favor of adopting the policy for good in the final period as a consequence of overoptimism. We find, however, very little evidence in support of persistent overoptimism for individual subjects in the lab implementation. By persistence we mean persistence within a round—that is, whether a subject who is overoptimistic after receiving an uncertain signal in the first period remains overoptimistic after another uncertain signal in the second period.

Some of the results may reflect the complexity of the strategic calculus for lab participants when best-response behavior given the actual behavior of other players conflict with equilibrium predictions. Also, as a caveat, the impact of individual mistakes on group decision may be lower in larger groups. Finding ways to provide subjects with more feedback about the behavior of other subjects without stretching the design beyond what is reasonable, and using larger groups, seem avenues worth exploring.

## Appendix A. Proofs

### *Proof of Proposition 2*

We proceed by backward induction. Since  $h > s > r$ , it is optimal for a player to vote for the risky project in the third period if the player received a high signal and against otherwise. Thus, in the third period, a group would decide to implement the project if and only if  $w_2 \geq 2$ .

Let us consider the voting in the second period. If a player receives a high signal, then it is a weakly dominant strategy to vote for the project since  $h > s$ . Hence, if there are at least two high signals received, then the majority would decide to continue the project. We are left to consider cases of  $w_1 = 0$  and  $w_1 = 1$ . For simplicity of exposition we consider these two cases as separate lemmas.

**Lemma 1.** *If  $w_1 = 0$ , voters will continue the policy in the second period if and only if*

$$\frac{h - s}{s - r} \geq \frac{1 - p_1q}{2 - p_1q}.$$

*Proof.* Suppose  $w_1 = 0$  and consider any of the three players. If the experimentation with the risky policy continues, it will be adopted in the final period in two cases:

- (i) The player obtains a high signal and at least another player does. The probability of this event is  $p_1q(2p_1q(1 - p_1q) + (p_1q)^2) = (p_1q)^2(2 - p_1q)$ , and in this case the player obtains a payoff of  $h$ .
- (ii) The player does not obtain a high signal but the other two players do. The probability of this event is  $(1 - p_1q)(p_1q)^2$ , and in this case the player obtains a payoff of  $r$ .

In the remainder event the policy is not adopted in the final period and the player obtains a payoff of  $s$ . Hence, the net gain of experimenting with the risky policy is:

$$E\pi(\text{Continue}) - s = (p_1q)^2 ((h - s)(2 - p_1q) - (s - r)(1 - p_1q)),$$

which is negative if and only if  $(h - s)/(s - r) < (1 - p_1q)/(2 - p_1q)$ . Hence, the player prefers to abandon the risky policy in the second period if and only if this inequality holds.  $\square$

**Lemma 2.** *If  $w_1 = 1$ , voters will abandon the policy in the second period if and only if*

$$\frac{h - s}{s - r} < 1 - p_1q.$$

*Proof.* Suppose  $w_1 = 1$ , and consider the players who did not receive a high signal. Their expected payoff can be decomposed in the following three cases:

- (i) The player receives a high signal in the next period. Then the player's payoff is  $h$  for sure, since there are at least two agents with high signals and the project will be implemented.
- (ii) The player does not receive a high signal but the other player does. Then the player's payoff is  $r$  since there are two players receiving high signals and the policy will be implemented.
- (iii) None of the two players receive high signals in the next period. Then, the player's payoff is  $s$  since the project will be abandoned.

The expected payoff from voting for continue experimenting is then:

$$E\pi(\text{Continue}) = p_1qh + p_1q(1 - p_1q)r + (1 - p_1q)^2s.$$

The alternative is to abandon the project and receive the safe payoff. Hence, the net gain of adopting the policy is

$$E\pi(\text{Continue}) - s = p_1q((h - s) - (1 - p_1q)(s - r)),$$

which is negative if and only if  $(h - s)/(s - r) < 1 - p_1q$ . Hence, under the condition of the lemma, the two players who did not receive the high signal prefer to abandon the policy in the second period.  $\square$

Intuitively, both lemma 1 and lemma 2 establish thresholds for the experimentation gain ratio  $(h - s)/(s - r)$  such that if the ratio is above that threshold it is optimal to continue experimenting after no news ( $w_1 = 0$ ) and after partial good news ( $w_1 = 1$ ). The threshold for partial good news,  $1 - p_1q$ , is higher than the threshold for no news,  $(1 - p_1q)/(2 - p_1q)$ . That is, no news is better than partially good news in terms of experimentation.

Finally, we show that if the conditions of lemmas 1 and 2 are satisfied, then the group will decide to initiate experimenting with the risky project. Note that players are identical ex ante, so experimentation will receive unanimous support.

**Lemma 3.** *If*

$$\frac{1 - p_1q}{2 - p_1q} \leq \frac{h - s}{s - r} < 1 - p_1q,$$

*voters will start experimenting with the policy in the first period.*

*Proof.* First, let us find the probability of receiving a high payoff for sure in the last period if the reform is started. There are two possible cases: either the player receives a high signal at  $t = 1$  or the player does not receive a high signal at  $t = 1$  but receives a high signal at  $t = 2$ . The following table summarizes the probabilities of implementing the policy conditional on possible events in these two cases.

High signal received at $t$	Prob of signal	Group signals $w_1$ $w_2$		Conditional prob of $(w_1, w_2)$
1	$p_0q$	2	2 or 3	$2p_0q(1 - p_0q)$
1	$p_0q$	3	3	$(p_0q)^2$
2	$(1 - p_0q)p_1q$	0	2	$(1 - p_0q)^2 2(p_1q)(1 - p_1q)$
2	$(1 - p_0q)p_1q$	0	3	$(1 - p_0q)^2 (p_1q)^2$
2	$(1 - p_0q)p_1q$	2	3	$(p_0q)^2$

In the table we use the fact that the project is stopped in the first period if and only if there is only one high signal received, as implied by lemmas 1 and 2 and the inequalities in the statement of this lemma.

Now, let us calculate the probability of receiving a payoff of  $r$  if the reform is started. In this case player cannot receive a high signal, which happens with probability  $(1 - p_0q)(1 - p_1q)$ . Then, the project is implemented if either both other players receive high signals at  $t = 2$  (which happens with probability  $(p_0q)^2$ ) or neither of them receive a high signal at  $t = 2$  and both receive high signals at  $t = 3$  (which happens with probability  $(1 - p_0q)^2 (p_1q)^2$ ).

Then, the expected net gain of adopting the policy in the first period is:

$$\begin{aligned} E\pi(\text{Continue}) - s &= (h - s) [p_0q (2p_0q(1 - p_0q) + (p_0q)^2) \\ &\quad + (1 - p_0q)p_1q ((1 - p_0q)^2(2(p_1q)(1 - p_1q) + (p_1q)^2) + (p_0q)^2)] \\ &\quad - (s - r)(1 - p_0q)(1 - p_1q) [(p_0q)^2 + (1 - p_0q)^2(p_1q)^2]. \end{aligned}$$

Substituting and simplifying,

$$\begin{aligned} E\pi(\text{Continue}) - s &= (h - s)(p_0q)^2 [2 - p_0q^2 + (1 - q)^2(2 - 3p_0q + p_0q^2)] \\ &\quad - (s - r)(p_0q)^2(1 - 2p_0q + p_0q^2)(1 + (1 - q)^2). \end{aligned}$$

Thus, every player prefers to adopt the policy in the first period if

$$\frac{h - s}{s - r} \geq \frac{(p_0q)^2(1 - 2p_0q + p_0q^2)(1 + (1 - q)^2)}{(p_0q)^2 [2 - p_0q^2 + (1 - q)^2(2 - 3p_0q + p_0q^2)]}, \quad (\text{A.1})$$

or equivalently

$$\frac{h - s}{s - r} \geq \frac{(1 - p_0q)(1 - p_1q)(1 + (1 - q)^2)}{1 + (1 - p_0q)[(1 - q)^2 + (1 - p_1q)(1 + (1 - q)^2)]}.$$

We can rewrite this inequality as

$$1 + \frac{1}{1 - p_1q} + \frac{p_0q}{(1 - p_0q)(1 - p_1q)(1 + (1 - q)^2)} \geq \frac{s - r}{h - s},$$

while we can rewrite the inequality  $(1 - p_1q)/(2 - p_1q) \leq (h - s)/(s - r)$  in the statement of the lemma as

$$1 + \frac{1}{1 - p_1q} \geq \frac{s - r}{h - s},$$

which is more demanding. □

Proposition 2 follows from Lemmas 1, 2 and 3.

*Proof of Proposition 3*

We proceed by backward induction. Since  $h > s > r$ , it is optimal for a player to vote for the risky project in the third period if the player received a high signal and against otherwise. Thus, in the third period, the group would decide to implement the risky project if and only if  $w_2 \geq 2$ . Moreover, there is not voting in the second period, so it is only left to show that subjects would prefer to initiate experimenting with the risky project under the proposed voting rule.



**Lemma 4.** *Suppose*

$$\frac{1}{2} \leq \frac{h-s}{s-r} < 1 - p_1q.$$

*Then, players vote to initiate the project in the first period under the optimal rule.*

*Proof.* The calculation is similar to the one in the proof of Lemma 3. To the events leading to receiving a high payoff for sure in the last period, we have to add the events in which only one high signal is received in the first period. The following table summarizes these additional events:

High signal received at $t$	Prob of signal	Group signals $w_1$ $w_2$		Conditional prob of $(w_1, w_2)$
1	$p_0q$	1	2	$(1 - p_0q)^2 2p_1q(1 - p_1q)$
1	$p_0q$	1	3	$(1 - p_0q)^2 (p_1q)^2$
2	$(1 - p_0q)p_1q$	1	2 or 3	$2p_0q(1 - p_0q)$

Respect to the calculation in the proof of Lemma 3, equation A.1, the probability of receiving a payoff of  $h$  in rather than  $s$  in the third period increases in

$$p_0q(1 - p_0q)^2 (2p_1q(1 - p_1q) + (p_1q)^2) + (1 - p_0q)p_1q \cdot 2p_0q(1 - p_0q),$$

or equivalently,

$$p_0q(1 - p_0q)^2 (2p_1q(1 - p_1q) + (p_1q)^2 + 2p_1q) \equiv A.$$

Similarly, to the events leading to receiving a expected payoff of  $r$  in the last period, we have to add the event in which only one high signal is received by other players in the first period, with the other player receiving a high signal in the second period, which has probability  $2(p_0q)(1 - p_0q)p_1q$ . Thus, respect to the calculation in the proof of Lemma 3, the probability of receiving a payoff of  $r$  in rather than  $s$  in the third period increases in

$$(1 - p_0q)(1 - p_1q)2p_0q(1 - p_0q)p_1q,$$

or equivalently,

$$p_0q(1 - p_0q)^2(1 - p_1q)(2p_1q) \equiv B.$$

Following the steps of the proof of Lemma 3, the condition for players to vote for adopting the policy in the first period is then

$$\frac{h-s}{s-r} \geq \frac{(p_0q)^2(1 - 2p_0q + p_0q^2)(1 + (1 - q)^2) + B}{(p_0q)^2 [2 - p_0q^2 + (1 - q)^2(2 - 3p_0q + p_0q^2)] + A}.$$

It is easy to check that  $A > B$ , so if inequality A.1 is satisfied, the condition above is satisfied

as well. □

## References

- Albrech, J., Anderson, A., Vroman, S., 2010. Search by committee. *Journal of Economic Theory* 145 (1386-1407).
- Bolton, G. E., Ockenfels, A., 2000. Erc: A theory of equity, reciprocity, and competition. *American economic review* 90 (1), 166–193.
- Buchanan, J., Tullock, G., 1962. *The Calculus of Consent*. The University of Michigan Press.
- Callander, S., Hummel, P., 2014. Preemptive policy experimentation. *Econometrica* 82, 1509–1528.
- Chan, J., Lizzeri, A., Suen, W., Yariv, L., 2015. Deliberating collective decisions. *Review of Economic Studies*.
- Chen, D. L., Schonger, M., Wickens, C., 2016. otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9, 88–97.
- Dolgoplov, A., Freer, M., 2018. Revealed social preferences. Available at SSRN 3111494.
- Elbittar, A., Gomberg, A., Martinelli, C., Palfrey, T. R., forthcoming. Ignorance and bias in collective decisions. *Journal of Economic Behavior & Organization*.
- Erev, I., Roth, A. E., 1998. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American economic review*, 848–881.
- Fehr, E., Schmidt, K. M., 1999. A theory of fairness, competition, and cooperation. *The quarterly journal of economics* 114 (3), 817–868.
- Fernandez, R., Rodrik, D., 1991. Resistance to reform: Status quo bias in the presence of individual-specific uncertainty. *American Economic Review*, 1146–1155.
- Khromenkova, D., 2017. Collective experimentation with breakdowns and breakthroughs, university of Mannheim (unpublished).
- McKelvey, R. D., Palfrey, T. R., 1995. Quantal response equilibria for normal form games. *Games and Economic Behavior* 10 (1), 6–38.
- McKelvey, R. D., Palfrey, T. R., 1998. Quantal response equilibria for extensive form games. *Experimental Economics* 1 (1), 9–41.

- Messner, M., Polborn, M. K., 2012. The option to wait in collective decisions and optimal majority rules. *Journal of Public Economics* 96 (5), 524–540.
- Moldovanu, B., Shi, X., 2013. Specialization and partisanship in committee search. *Theoretical Economics* 8 (3), 751–774.
- Strulovici, B., 2010. Learning while voting: Determinants of collective experimentation. *Econometrica* 78 (3), 933–971.

## Appendix B. Estimation of expected earnings (For Online Publication Only)

In this section we present details on the expected and average welfare. We compute the expected earnings conditional on the realized path of group signals  $(w_1, w_2)$  even if subjects decide to abandon the risky project before learning the realization of the path. The second column in Table B.9 presents the theoretical probabilities of each given path, the third column the estimated probability for the majority treatment and last one for the optimal one. Types and signals have been randomly generated before subjects start voting, and we consider the entire possible path without conditioning on whether the group decided to start and/or continue the project. Probabilities in both treatments differ significantly from the theoretical predictions. Moreover, the differences between the treatments are also significant. Hence, it is important to control for the distribution of types and signals in order to perform an expected welfare comparison.

$(w_1, w_2)$	Theoretical	Majority	Optimal
(0,0)	.244	.213	.283
(0,1)	.147	.104	.142
(0,2)	.029	.017	.033
(0,3)	.002	.004	.000
(1,1)	.293	.304	.213
(1,2)	.117	.121	.146
(1,3)	.012	.015	.025
(2,2)	.117	.156	.142
(2,3)	.030	.058	.013
(3,3)	.016	.017	.004

Table B.9: Frequency of signals observed by treatment

For the purpose of welfare analysis, we conduct a two-step procedure to estimate expected revenues. The algorithm selects paths according to the theoretical distribution from Table B.9. Then, it randomly picks observations from the subsample of observed decisions given the history of the signals chosen. Bootstrapped sample sizes correspond to the observed sample sizes. Further, the algorithm calculates expected revenues (given the theoretical distribution), and allows us to create the distribution of expected earnings for every treatment. We generate a sample of 10000 expected earnings using this technique. The results are based on these estimates.

Figure B.7 presents expected earnings obtained from the procedure explained above. The bootstrapped means are  $-3.26$  (median is  $-3.11$ ) for majority and  $-14.53$  (median is  $-14.26$ ) for the optimal treatment. Moreover, the expected earnings from the optimal treatment are significantly below those from the majority treatment ( $p < .001$ ). In addition we compute the expected welfare with the pooled data (both majority and optimal treatment) and obtain

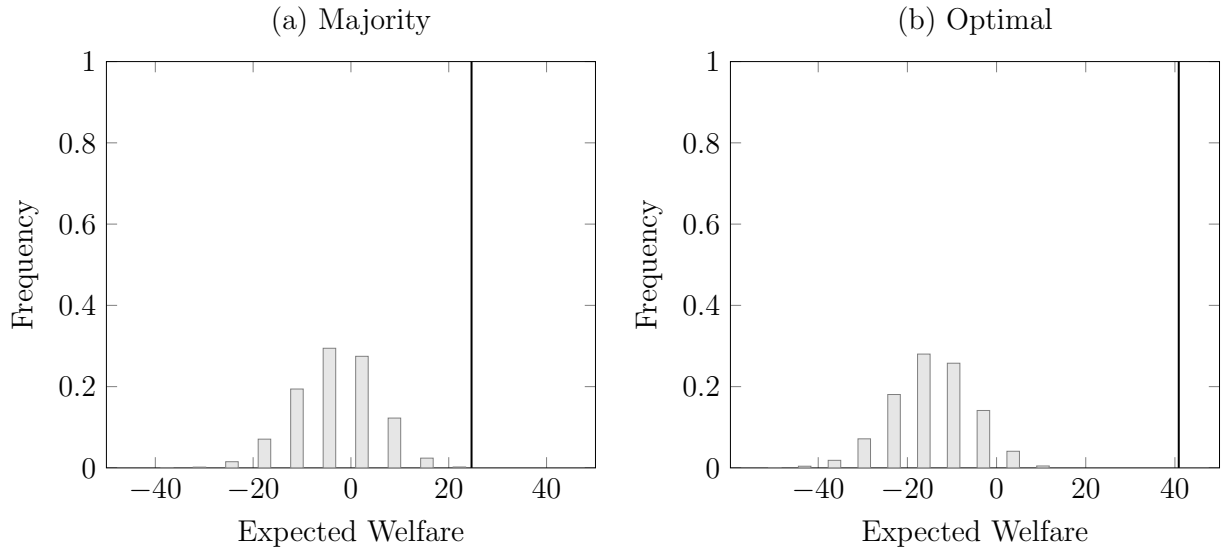


Figure B.7: Expected Earnings

the expected welfare is  $-8.78$  (median is  $-8.37$ ). Moreover, the merged expected welfare is significantly ( $p < .0001$ ) lower than the one from majority treatment and significantly ( $p < .0001$ ) higher than the one from optimal treatment. Vertical lines show the equilibrium level of earnings predicted by the theoretical model. We see that both treatments lie significantly below the equilibrium earning levels.

## Appendix C. Robustness (For Online Publication Only)

In this Appendix we present the robustness checks, tables and graphs mentioned in the paper. First, we present the robustness checks on learning behavior and Next, we present the robustness check on the welfare estimates for the structural model given different quantal response coefficients.

### Appendix C.1. Learning

We explore this possibility using a simple regression specification  $y_\tau = \beta\tau + c$ , where  $y_t$  is the variable of interest,  $\tau$  is the round and  $c$  is a constant. Further we present the results of the regression analysis of equilibrium behavior and voting for the risky project given the signal for all cases considered above.

	Majority						Optimal		
	$t = 2$			$t = 3$			$t = 3$		
	$s_i = 0$	$s_i = 1$	any	$s_i = 0$	$s_i = 1$	any	$s_i = 0$	$s_i = 1$	any
$w_t = 0$	.3330	–	.3329	.6330	–	.6330	.0331	–	.0331
$w_t = 1$	.8889	.5010	.8038	.0004	.9238	.0020	.0343	.0891	.3457
$w_t \geq 2$	.3154	.8037	.5980	.3371	.2025	.9500	.6970	.4521	.6718
any	.6773	.5612	.5216	.2799	.3507	.0216	.0064	.9965	.0512

Table C.10: Results of regression analysis of effect of learning over the frequency of equilibrium action. We report  $p$ -values according to the probit specification in the cells of the table.

Table C.10 presents the  $p$ -values for learning regressions decomposed by the signal received by subject and the total amount of high signals received by the group. Directions are the same as in Table 4. The only reasonable difference appears at the case of  $w_2 = 1$  and  $s_2 = 1$  in the optimal treatment, where the learning is significant at 10% level according to Table C.10, but is not significant according to Table 4.

### Appendix C.2. Structural Model

We use the structural model with parameters estimated using majority treatment, optimal treatment, and the pooled data to simulate expected social welfare under the majority and under the optimal treatment.

	Majority	Optimal
Majority treatment estimates	9.43	-76.55
Optimal treatment estimates	13.92	-0.53
Pooled data estimates	14.32	-2.28

Table C.11: Expected welfare for the estimated coefficients

Table C.11 presents the results. Estimated welfare for the majority treatment is higher for the optimal rule treatment. Note that the estimated fraction of overoptimistic people is larger in the first row than in the second row. Increasing the fraction of overoptimistic subjects decreases welfare under both treatments. As is illustrated in Table C.11), the optimal mechanism is more vulnerable to such a change.

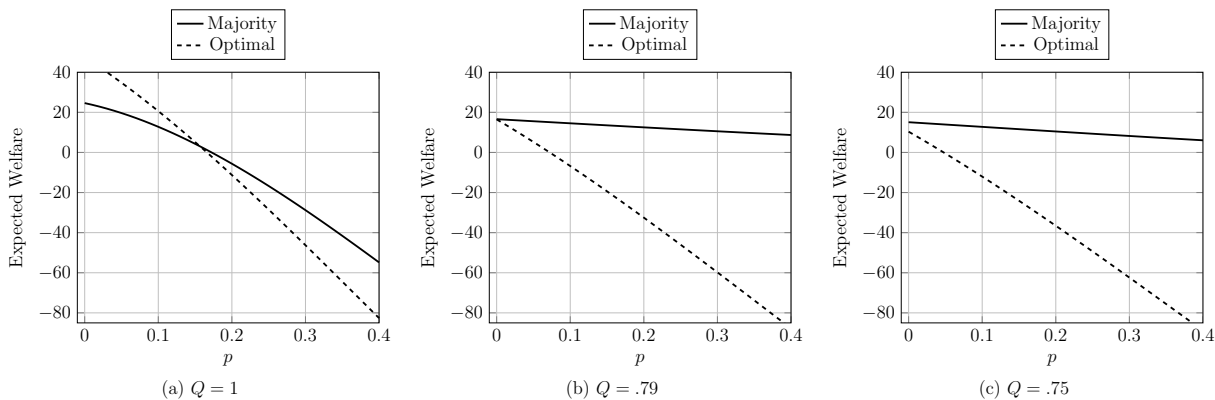


Figure C.8: Expected welfare estimated using the structural model for different levels of overoptimism given alternative estimates of  $Q$ .

Figure C.8 presents the expected welfare using the quatal response model with overoptimistic subjects at different levels of over-optimism. Figures C.8(b) and C.8(c) present the welfare for the levels of  $Q$  estimated using the Majority treatment and Optimal treatment data separately. In both cases the expected welfare from the majority mechanism is higher than expected welfare from the optimal mechanism for all levels of overoptimism. Moreover, the difference is increasing as the level of overoptimism increases.

Note that the higher is  $Q$  the higher would be the expected welfare (*ceteris paribus*) delivered by the optimal mechanism. Moreover, the higher is overoptimism level given fixed  $Q$  the lower would be the expected welfare delivered by optimal mechanism. Figure C.8(a) presents the expected welfare for different overoptimism levels given  $Q = 1$ . For the low enough levels of overoptimism ( $p \leq .15$ ) optimal mechanism delivers higher welfare than the majority one. However, for the higher levels of overoptimism majority mechanism becomes more profitable.

Another dimension of comparisons shows the robustness of the majority treatment. Moreover, for the imperfectly rational players ( $Q < 1$ ) the declining curve for majority mechanism becomes flatter. Hence, the marginal welfare loss from overoptimism is lower for imperfectly rational players. At the same time there this trend is not present for the optimal mechanism and slope of the curve is pretty much independent on  $Q$ .

*Appendix C.3. Structural Model with non-persistent over-optimism*

One of the clear observations which comes out of our analysis is that over-optimism is rather non-persistent. That is subjects in the second period are rather more over-optimistic than in the last round. Hence, we use less restrictive model which allows for non-persistent over-optimism. That is the probability of over-optimistic behavior depends on the stage of the game. Hence, instead of having unique over-optimism parameter there are three different coefficients  $p_1, p_2, p_3$  for the over-optimism in each stage of the game.

Regardless the fact that the model is more realistic and more flexible we decided to exclude from the main text, since it produces several artifacts which we do not find realistic. In particular estimating the model using the Optimal treatment data implies that the best-response in the first stage is not to start the experimentation. Given that such prediction contradicts the observations, hence, the probability of over-optimism at the first-stage is almost one.

We find such artifact to be unrealistic, hence, we also estimate the restricted model using Optimal treatment data. That is we restrict the parameters such that it is still a best-reply to start the experimentation. We refer to this model as Constrained Optimal. Before we proceed let us note that we obtained that the models are significantly better explain the data ( $p < .001$ ) for all cases. Hence, these models explain the data significantly better.

	$Q$	$p_1$	$p_2$	$p_3$
Majority treatment	.79	.1	.56	.17
Optimal treatment	.78	.91	–	.23
Constrained Optimal treatment	.85	.05	–	.07

Table C.12: Estimated parameters

Table C.12 presents the estimates for the coefficients for all three cases. In this case we abstained from estimation on the pooled data the models are not compatible. Parameters for Majority and Optimal treatment are close once we consider  $p_3$  (probability of over-optimistic behavior in stage 3) and  $Q$  (probability of best-responding). Moreover, probability of best-responding ( $Q$ ) is close to that of original estimation. However, as we mentioned above, the probability of over-optimistic behavior in the stage one is almost one for Optimal treatment. Once we consider the Constrained model for Optimal treatment, the probability of best-responding is much higher and the levels of over-optimism are rather close to those of original estimation for the Optimal and Pooled data.

Table C.13 presents the conditional implementation probabilities and the expected welfare. By conditional implementation probabilities we mean the probability that reform is implemented given  $w_2$  high signals received before that stage 3 as in Table 2. First, let us note that numbers for Majority and Optimal treatments are quite close to the corre-



	$w_2 = 0$	$w_2 = 1$	$w_2 = 2$	$w_2 = 3$	Expected Welfare
Majority treatment	.09	.27	.67	.85	-14.36
Optimal treatment	.18	.44	.89	.98	-35.67
Constrained Optimal treatment	.07	.27	.82	.97	$\approx 0$

Table C.13: Conditional implementation probabilities and expected welfare

sponding numbers from Table 2. That is consistent with the fact that these models describe the data quite well. We can see that Optimal treatment estimate captures the significant over-implementation observed in the data. At the same time if we consider the Constrained Optimal estimate, then the deviation comes in particular in the frequency of the implemented projects if no high signals received. Given the reduced form evidence we have, this frequency should be higher for the majority treatment, but coefficients from the Constrained Optimal estimate show the reverse order.

Finally, we compare the expected welfare from different mechanisms. We use the estimates from the corresponding models. Expected welfare from the Majority mechanism is higher than one of Optimal mechanism. However, if we use the Constrained optimal estimate for the Optimal mechanism, then the expected welfare from the Optimal mechanism is higher than the expected welfare from the Majority mechanism. Hence, we can obtain different welfare implications depending on whether we trust that Constrained Optimal model is rather real one or not.

## **Appendix D. Experimental instructions (For Online Publication Only)**

We provide first the instructions for the majority treatment and next the instructions for the optimal voting rule treatment.

## Instructions

Welcome to our experiment! You have earned \$5 by showing up on time. If you read and follow the instructions below carefully, you have the potential to earn up to \$35. In the experiment you will earn Experimental Dollars (E\$) which will be converted into cash at the end of the experiment. **For every 20 E\$ you have at the end of the experiment you will be paid 1 US Dollar.** You will NOT be told the names of those in your group and they will NOT be told your name. All participants have identical instructions. NO communication with other participants is allowed during this experiment. Please switch off your cell phones. If you have any questions please raise your hand, and the experimenter will assist you individually.

## Decision Task

There are 15 rounds in this experiment. At the end of this experiment, only 1 round will be randomly selected to determine your final payment. Every round has equal chance of being chosen as the payoff round, therefore **it is in your best interest to treat each round as if it is the one that determines your payment.**

At the beginning of each round, three players will compose a group, and each of you will get a Player ID (for example, Player 1, Player 2 or Player 3). **The Player ID may vary from round to round. For each round, you will be randomly placed in a new group.**

There is one project at each round (different project at different round). You and the other two group members jointly decide whether or not to implement a project. If a project is implemented, each of you has 50% chance to receive a high payoff: **E\$500**, and a 50% chance to receive a low payoff: **E\$50**. Whether each of you receive a high payoff or low payoff is independently and randomly determined at the beginning of each round. If the project is NOT implemented, each of you will receive **E\$350** for sure.

There are **three stages** for each round. At each stage, each of you will choose an action. The **majority action of the group** (i.e. at least two players) will be the action realized.

At Stage 1, each of you can choose **INITIATE** or **STOP** the project right away. If the majority of you choose to **STOP**, the project is NOT implemented and this round ends. If the majority of you choose to **INITIATE** the project, each of you will get a signal regarding whether you can receive high payoff if the project is implemented. The signal works in the following way:

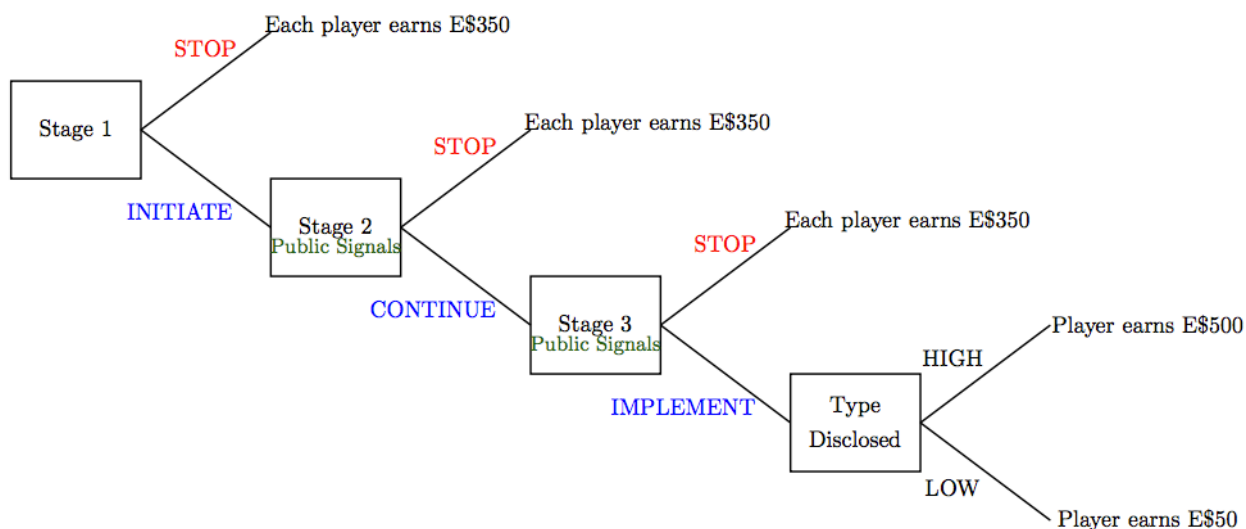
The signal is programmed to be either “**High payoff for sure**” or “**Uncertain.**” If you are a player who will receive **high payoff** if the project is implemented, **there is 50% chance for you to receive a signal that says “High payoff for sure”, and another 50% to receive a signal that says “Uncertain.”** If you are a player who will receive the **low payoff** if the project is implemented, **you will always receive a signal that says “Uncertain.”** In other words, if you get a “**High payoff for sure**” signal, you will keep receiving the same signal in the following

stage (if any) of the same round, and you know that this round you will receive a high payoff for sure if the project is implemented. If you receive a signal that says “**Uncertain**”, it’s not clear whether you will receive high or low payoff if the project is implemented. Your signal only reveals your own chance of receiving high or low payoff from this project. Your group members receive different signals regarding their chance. All the signals are public information, that is, they will be observed by all group members.

At Stage 2, you can choose to **CONTINUE** the project or **STOP**. If the majority of you choose to **STOP**, the project is NOT implemented and the round ends. If the majority of you choose to **CONTINUE** the project, each of you will get another signal about whether you are a high payoff player. The signal works exactly the same way as we explained above.

At Stage 3, you can choose to **IMPLEMENT** the project or **STOP**. If the majority of you choose to **STOP**, the project is NOT implemented and this round ends. Otherwise, the project will be implemented. Each group member’s payoff will be revealed at the end of the round.

The following graph summarizes the possible decisions and outcomes at each stage. Please raise your hand if you have any questions.



**Thank You and Good Luck!**

## Instructions

Welcome to our experiment! You have earned \$5 by showing up on time. If you read and follow the instructions below carefully, you have the potential to earn up to \$35. In the experiment you will earn Experimental Dollars (E\$) which will be converted into cash at the end of the experiment. **For every 20 E\$ you have at the end of the experiment you will be paid 1 US Dollar.** You will NOT be told the names of those in your group and they will NOT be told your name. All participants have identical instructions. NO communication with other participants is allowed during this experiment. Please switch off your cell phones. If you have any questions please raise your hand, and the experimenter will assist you.

## Decision Task

There are 30 rounds in this experiment. At the end of this experiment, only 1 round will be randomly selected to determine your final payment. Every round has equal chance of being chosen as the payoff round, therefore **it is in your best interest to treat each round as if it is the one that determines your payment.**

At the beginning of each round, three players will compose a group, and each of you will get a Player ID (for example, Player 1, Player 2 or Player 3). **The Player ID may vary from round to round. For each round, you will be randomly placed in a new group.**

There is one project at each round (different project at different round). You and the other two group members jointly decide whether or not to implement a project. If a project is implemented, each of you has 50% chance to receive a high payoff: **E\$500**, and a 50% chance to receive a low payoff: **E\$50**. Whether each of you receive a high payoff or low payoff is independently and randomly determined at the beginning of each round. If the project is NOT implemented, each of you will receive **E\$350** for sure.

There are **three stages** for each round. At each stage, each of you will choose an action. The **majority action of the group** (i.e. at least two players) will be the action realized.

At Stage 1, each of you can choose **INITIATE** or **STOP** the project right away. If the majority of you choose to **STOP**, the project is NOT implemented and this round ends. If the majority of you choose to **INITIATE** the project, each of you will get a signal regarding whether you can receive high payoff if the project is implemented. The signal works in the following way:

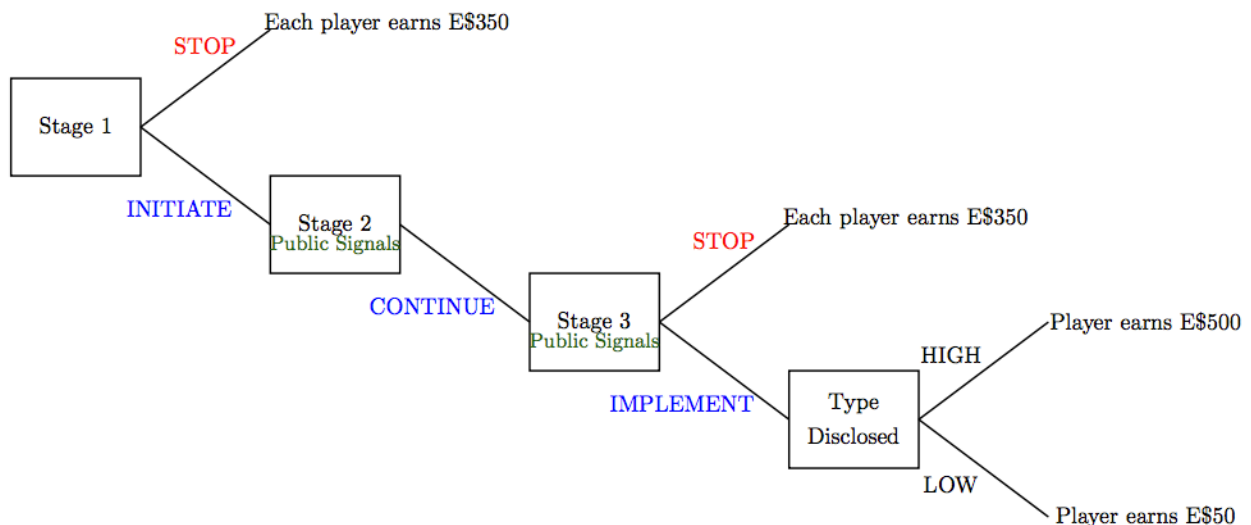
The signal is programmed to be either “**High payoff for sure**” or “**Uncertain.**” If you are a player who will receive **high payoff** if the project is implemented, **there is 50% chance for you to receive a signal that says “High payoff for sure”, and another 50% to receive a signal that says “Uncertain.”** If you are a player who will receive the **low payoff** if the project is implemented, **you will always receive a signal that says “Uncertain.”** In other words, if you get a “**High payoff for sure**” signal, you will keep receiving the same signal in the following

stage (if any) of the same round, and you know that this round you will receive a high payoff for sure if the project is implemented. If you receive a signal that says “**Uncertain**”, it’s not clear whether you will receive high or low payoff if the project is implemented. Your signal only reveals your own chance of receiving high or low payoff from this project. Your group members receive different signals regarding their chance. All the signals are public information, that is, they will be observed by all group members.

At Stage 2, the project continues. Each of you will get another signal about whether you are a high payoff player. The signal works exactly the same way as Stage 1.

At Stage 3, you can choose to **IMPLEMENT** the project or **STOP**. If the majority of you choose to **STOP**, the project is NOT implemented and this round ends. Otherwise, the project will be implemented. Each group member’s payoff will be revealed at the end of the round.

The following graph summarizes the possible decisions and outcomes at each stage. Please raise your hand if you have any questions.



**Thank You and Good Luck!**