FEDERAL
RESERVE
BANK
of ATLANTA

**Corporate Board Composition, Protocols, and Voting Behavior: Experimental Evidence**

Ann B. Gillette, Thomas H. Noe, and Michael J. Rebello

Working Paper 2000-10
August 2000

# Working Paper Series

# Corporate Board Composition, Protocols, and Voting Behavior: Experimental Evidence

Ann B. Gillette, Thomas H. Noe, and Michael J. Rebello

Federal Reserve Bank of Atlanta
Working Paper 2000-10
August 2000

**Abstract:** We model experimentally the governance of an institution. The optimal management of this institution depends on the information possessed by insiders. However, insiders, whose interests are not aligned with the interests of the institution, may choose to use their information to further personal rather than institutional ends. Researchers (e.g., Palfrey 1990) and the business press have both argued that multiagent mechanisms, which inject trustworthy but uninformed "watchdog" agents into the governance process and impose penalties for conflicting recommendations, can implement institutionally preferred outcomes. Our laboratory experiments strongly support this conclusion. In the experimental treatments in which watchdog agents were included, the intuitionally preferred allocation was implemented in the vast majority of cases. Surprisingly, implementation occurred even in the absence of penalties for conflicting recommendations.

JEL classification: G3, C7

Key words: corporate governance, implementation, experimental economics, mechanism design

---

The authors thank participants of the Atlanta Finance Workshop and the Economic Science Association Meetings for useful comments. Gillette acknowledges research support from the College of Business Administration, Georgia State University. This paper was completed while Gillette was a visiting scholar at the Federal Reserve Bank of Atlanta. The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta or the Federal Reserve System. Any remaining errors are the authors' responsibility.

Please address questions regarding content to Ann B. Gillette, Department of Finance, J. Mack Robinson College of Business, Georgia State University, University Plaza, Atlanta, Georgia 30303-3083, 404-651-2671, 404-651-2630 (fax), agillette@gsu.edu; Thomas H. Noe, A.B. Freeman School of Business, Tulane University, New Orleans, Louisiana 70118-5669, 504-865-5425, tnoe@mailhost.tcs.tulane.edu; or Michael J. Rebello, Department of Finance, J. Mack Robinson College of Business, Georgia State University, University Plaza, Atlanta, Georgia 30303-3083, 404-651-2781, mrebello@gsu.edu.

# Corporate Board Composition, Protocols, and Voting Behavior: Experimental Evidence

## 1 Introduction

Consider the situation of owners of a corporation when they entrust the fate of their institution to groups of agents, henceforth called "insiders." The insiders' interests may not be fully aligned with those of the owners, and the owner-preferred allocation may be contingent on private information possessed by the insiders. In this situation, owners need a mechanism by which to mitigate the the harm caused by conflicts of interest. Such a mechanism may determine both the composition of the group of insiders and the environment within which the group functions. The same need arises in the public sector when citizens elect representatives or governments delegate authority over public enterprises to boards.

The theory of mechanism design provides a great deal of insight into the design and effectiveness of boards. Mechanisms that implement the institutionally preferred policy are fairly easy to devise. However, such mechanisms may implement inefficient policies as well. It may also be possible to devise mechanisms that fully implement the institutionally preferred policy, that is, mechanisms whose only outcome is the institutionally preferred policy (see, e.g., Palfrey and Srivastava, 1993). These mechanisms display two features observed in "real-world" corporate and civic boards. First, penalties are imposed when insiders dissent from the board consensus.[1] Second, institutions are designed to foster the participation of uninformed outside "watchdog" agents whose interests are aligned with the principal.[2]

---

[1] Punishments for dissident members of organizations are frequently observed in a number of real-world institutions (Warther (1998)). For example, rarely do dissident directors survive on corporate boards. A particularly notable case is where the truthful dissident messenger, the former CFO and board member of Apple Computer Inc., Joseph Graziano, was "shot-down" by the board after placing strong objections to the CEO's yearly business plan, and promptly resigned the next day.

[2] The inclusion of uninformed, watch-dog agents in institutional governance is common place.

Penalties for disagreement and the presence of outside agents are both important for implementing institutionally preferred outcomes. Because insiders receive the same information regarding the payoffs from policies, insiders can forecast the recommendations of other insiders. Thus, if their is a penalty for disagreement and if no single insider can sway the outcome of the decision with his recommendation, supporting a policy that is harmful to the institution is costly to any one insider if all other agents are acting in the institution's interest. Disagreement penalties ensure that there exist equilbiria in undominated strategies that support the institutionally preferred outcome. However, absent outside agents, their may also exist equilibira that support undesirable outcomes. Full implementation of institutionally preferred policies results from the introduction of uninformed "watch-dog" agents whose interests are aligned with the institution. Such agents, though uninformed, have rational expectations; thus, thus they correctly conjecture the probablity distribution for project choice that would be generated if insiders were allowed to determine project choice. Thus watchdogs have an incentive to veto egregious policy choices by insider and, the veto

---

Examples include, ouside directors sitting on corporate boards, citizen review boards for police departments, holding referenda for bills approved by legislators. Inclusion of outsiders in the decision making process is lauded in the popular press. For example, the December 1997 issue of Business Week's special report on corporate governance, "The Best and Worst Boards," stated, "Another crucial component of top-ranked boards is director independence. Governance experts believe that a majority of directors should be free of all ties to either the CEO or the company. That means a minimum of insiders on the board, with directors and their firms barred from doing consulting, legal, or other work for the company. Interlocking directorships - in which CEO's serve on each other's boards - are also out ... But perhaps the best guarantee that directors act in shareholders' best interests is the simplest: Most good boards now insist that directors own significant stock in the company they oversee." Independent watchdog participation also appears to be on the increase, judging by the changes wrought in corporate board membership by shareholder activists. One of a growing number of shareholder activist victories in attaining more outside directors and stricter definitions for outside directorships occurred in 1996 at Archer Daniels Midland.

threat vitiates undesirable equilibria (Palfrey, 1990).

However, multi-agent mechanisms may still fail to fully implement institutionally preferred policies for a number of reasons. First, coordination between agents may fail and convergence to a Nash equilibrium may not occur. Second, coordination between insiders may be "too successful," leaving the equilibria that support the institutionally preferred allocation susceptible to the resulting coalition formation. In other words, the equilibria may not be "coalition-proof" (Bernheim, Peleg, and Whinston, 1987): exist "self-enforcing"strategy vectors involving simultaneous deviations by a subset of agents may produce a higher payoff to that subset.[3] Whether or not these difficulties will prevent the implementation of institutionally preferred allocations may depend in a rather subtle way on how communications are structured between agents (Milgrom and Roberts, 1996) and on agents' conjectures regarding the efficacy of cheap talk (Farrell and Rabin, 1996).

A priori, it is difficult to determine how these issues are resolved in real economic mechanisms. While a considerable amount of research has focused on the composition and performance of corporate boards, scepticism regarding board effectiveness remains (see, e.g., Jensen, 1993).[4] In some measure, this continuing debate on the effectiveness of boards results from difficulties in measuring the day-to-day effect of board composition on corporate performance (e.g., see Hermalin and Weisbach (1991)).

[3] In this context, self-enforcing strategy vectors are ones where, holding fixed the strategy set of the non-deviating agents, the deviating agents do not have an incentive to "double-cross" other deviating agents by defecting from the deviating coalition.

[4] Empirical studies on this subject tend to focus on the impact of board composition, as measured by the proportion of inside directors (management) and outsiders (those not otherwise affiliated with the firm). In particular, Weisbach (1988) finds that CEO turnover is greater and firm performance is superior when there is a larger fraction of outside directors. Furthermore, board composition appears to matter in various types of hostile and non-hostile takeover attempts (see Brickley et al (1994), Byrd and Hickman (1992), and Shivdasani (1993)).

To assess mechanisms designed to implement institutionally preferred policies, we simulate a board in a laboratory experiment with human subjects. Our subjects play a game that embeds several features of the multi-agent problem discussed above. The subjects are divided into two groups–a group of insiders and a group of watchdogs. Together, they must decide whether to accept a project. The project affects the value of the institution. Watchdog agents' payoffs are aligned with the institution's value. Insiders have private information that indicates whether the investment project increases value. However, their objective function is such that they gain from undertaking the project whether or not institutional value increases. A vote by the agents determines whether the project is undertaken. Split insider votes (some yes, some no) engender a positive probability that insiders will be penalized.

The game has both "efficient" outcome and an "inefficient" equilibirum outcomes. The efficient outcome is supported by many equilibria. Under the efficient outcome, the project is accepted only when it is value increasing and insider votes are not split.[5] In these efficient equilibria, insiders vote in favor of the project only when it is value increasing and watchdog agents abstain from voting. The second outcome, the inefficient outcome, is for the project to be rejected regardless of its quality. Equilibria supporting this outcome call for watchdogs to block acceptance of the project. The inefficient outcome is strictly dominated by the efficient outcome in the sense that all agents payoffs are strictly higher under the efficient outcome. Despite this dominance relationship, the equilibria supporting the efficient outcome are not coalition-proof (Bernheim, Peleg, and Whinston (1987)). All coalition-proof equilibria support the inefficient outcome.

The results of our experiments provide strong evidence that multi-agent voting mechanisms can implement efficient policies. The mechanism produced the efficient outcome the vast majority of times, even in our base case, when the communication protocol would should have maximized the incentive for coalitional defection: First all agents engaged in non-binding preplay communication; then, unobserved by watchdogs, insiders engaged in preplay communication amongst themselves.

[5] Note that we will use the term "efficient policy" to describe outcomes in which only the first of these conditions need be satisfied.

Treatments featuring other communication protocols produced statistically indistinguishable results.

In summary, multi-agent mechanisms of the type proposed by Palfrey (1990) can implement efficient policies for a wide range of agent communication protocols even in the presence of coalitional defection incentives. These results encourage organizational designs that give watchdog agents majority voting power in organizations; they justify recent tendencies in corporate governance to give outside shareholders majority voting rights and tendencies in civil governance to give voters veto power over legislative initiatives through referenda.

Within the experimental economics literature, our analysis lies at the interstices of research on communication, voting, and implementation. The communication literature led us to consider the effects of a wide variety of communication protocols.[6] However, because our concern is the viability of multi-agent mechanisms not communication per se, communication is treated a a control variable rather than an object of study. Similarly, though we are concerned with how strategic behavior of agents affects the outcomes of voting mechanisms, our focus diverges from the focus of the voting literature. This literature focus on agenda-setting, (s, e.g., Eckel and Holt, 1989). We hold the agenda for our voting game fixed while varying other factors such as the communication opportunities and cost of disagreement.

[6] Cooper et al.(1989) found that nonbinding preplay communication was a way to resolve coordination problems. However, failure to coordinate at the communication stage made the second stage more noncooperative. They show that two-sided, nonbinding communication reduces but does not eliminate coordination problems in the battle-of-the-sexes game, while one-sided communication of a nonbinding preplay message can avoid the possibility of crossed intentions, and largely eliminate coordination failure. Isaac and Walker (1988) note that experimental studies asking subjects make binary decisions between cooperation and defection find general support for the premise that communication significantly improves cooperation. See, Farrell and Rabin (1996) for a general review of these issues.

The experimental literature on implementation explores the viability of implementation mechanisms. However, this literature focuses on implementation with two agents in games of complete information, such as Abreu-Matsushima mechanisms (Sefton and Yavas, 1996). In contrast, we consider implementation in settings characterized by many agents and incomplete information. Thus, unlike two-agent settings, our experimental design makes coalition formation, and the format in which communication takes place, central.

The remainder of the paper is organized as follows. Section 2 describes the game, delineates the equilibria, and analyzes their properties. Section 3 describes the experimental design. Section 4 presents our results. Section 5 concludes the paper and suggests directions for future research.

## 2 Model

In this section, we present our basic model and characterize the coalition-proof outcomes and efficient Nash outcomes. The results of this section highlight the basic tension in our model—no efficient outcome is coalition-proof and no coalition-proof outcome is efficient. In fact, the Nash efficient outcome is preferred by all agents to the coalition-proof outcome.

At an intuitive level, the model is straightforward. The institution must choose whether to accept or reject a project. Under some insider information signals, project acceptance is good for the institution and under other signals it is bad. On average, acceptance is bad for the institution. However, acceptance always benefits insiders. These informed insiders each make a recommendation to outsiders. Subsequently, both insiders and outsiders vote. Outsiders, whose interests are aligned with the institution, have voting majority.

Insiders can either recommend honestly—that is, recommend the project only when it is in the institution's best interest—or recommend deceptively—that is always recommend acceptance. Because of penalties for dissent, a lone insider has no incentive to deviate to deceptive recommendations given that other insiders are making honest recommendations. If the informed insiders are making honest recommendations, then uninformed outsiders have an incentive to go along. Thus, the institutionally preferred outcome can be implemented by a Nash equilibrium. However, if outsiders "go along" with insider recommendations, then insiders have an incentive, as a coalition, to agree to provide misleading recommendations that ensure project acceptance in all states. Rational outsiders, anticipating insider collusion, will then ignore the insiders' recommendations and block project acceptance, even though they know that, under some insider information signals, accepting the project is good for the institution. Thus, the coalition-proof outcome is universal rejection, an outcome that neither insiders or outsiders desire.

Formalizing this argument is tedious for two reasons. First, representing communication over a general message space requires much notation. Second, the standard definition of coalition-proof Nash equilibria is inductive. Thus, all formal proofs relating to coalition proofness, even the most trivial, must be established using the machinery of mathematical induction. However, in the interest of rigor, we provide a formal analysis of the equilibria below.

*2.1 Agents*

The game is played by $[N] = \{1, 2, \dots N\}$ agents. The first $W$ agents are "watchdog" agents. The set of watchdog agents, which we represent by $[W]$, is given by $[W] = \{1, 2 \dots W\}$. The next $I$ agents are "insider" agents. The set of insider agents is given by $[I] = \{W + 1, W + I, \dots N\}$. We assume that $W > I > 1$.

*2.2 Information*

Before voting, insiders receive an information signal, $s$, revealing the quality of the project. This signal can take on one of two possible values: good, $G$, or bad, $B$. Watchdogs have a probability distribution over $G$ and $B$ which places weight $\pi$ on outcome $G$ and weight $1 - \pi$ on outcome $B$.

*2.3 Actions and Strategies*

*2.3.1 Communication Stage*

The game consists of two stages: a communication stage and a decision stage. All communication is "cheap" in that communication has no direct effect on the welfare of agents. A communication strategy for the uninformed watchdog agents is a message from a message space $M$. Let $\mathbf{m}_i^W \in M$, represent the message strategy of the $i$th watchdog. For the informed insider agents, a message strategy is a map from the information signal space $\{G, B\}$ into the message space $M$. For each insider, let $\mathbf{m}_i^I \in M^{\{G,B\}}$ represent the insider's communication strategy.

*2.3.2 Voting stage.*

Conditioned on the information signal sent in the first stage, the agents vote on whether to undertaking the project. Insider agents either vote in favor of the project, $\mathcal{Y}$, or against the project, $\mathcal{N}$. Watchdogs agents either vote against the project, $\mathcal{N}$, or abstain, $\mathcal{A}$.[7] Let $v = (v_1, \ldots v_N)$ represent the vector of votes by the $N$ agents and let $V$ represent the set of all possible vote vectors. Let $v^W$ represent the vector consisting of the first $W$ components, the watchdog votes, of the vector $v$ and let $v^I$ represent the subvector consisting of the last $I$ components, the insider votes. For any vector (or subvector of), $v$ let $\#\mathcal{Y}(v)$ represent the number of yes votes and let $\#\mathcal{N}(v)$ represent the number of no votes in a vector $v$.

The communication stage produces a message from each agent. Thus, a voting strategy for an individual agent is a map from observed messages into votes. Let $M^N$ represent a vector of messages sent by the agents. A voting strategy for watchdogs, which we represent by $\mathbf{v}_i^W$, is a map from $M^N$ to $\{\mathcal{N}, \mathcal{A}\}$. A voting strategy for insiders, $\mathbf{v}_i^I$, is a map from $\{G, B\} \times M^N$ to $\{\mathcal{Y}, \mathcal{N}\}$.

*2.3.3 Strategies*

A strategy for an individual agent is thus an ordered pair of communication and voting strategies. We represent a strategy of watchdog agent $i$, by $\sigma_i^W \equiv (\mathbf{m}_i^W, \mathbf{v}_i^W)$ and the strategy of an insider agent $i$, by $\sigma_i^I \equiv (\mathbf{m}_i^I, \mathbf{v}_i^I)$.

Note that the message and voting strategies define an overall strategy vector for insiders and watchdogs. This strategy for insiders can be viewed as a map from the information signal $s$ into

[7] Theoretically, it does not matter whether the voting space is "yes", "no", and "abstain" for both types of traders, versus the restricted voting space described above. Restricting the watchdogs' space to "abstain" or "no" helps us discern whether the watchdogs are voting with the recommendation of the insiders, and reduces the complexity of the coordination problem faced by the watchdogs.

$M \times \{Y, N\}^{M^N}$, given by

$$s \xrightarrow{\sigma_i^I} (\mathbf{m}^I(s), \mathbf{v}^I(s, m)).$$

A strategy for a watchdog, $\sigma_i^W$, is simply an element of $M \times \{Y, N\}^{M^N}$. Let $\sigma^I$ represent the vector of insider strategies; let $\sigma^W$ represent the vector of watchdog strategies.

Strategies yield votes via the following functional compositions. The message strategies of the insiders and watchdogs, and the information signal, $s \in \{G, B\}$, will produce a pattern of messages. These messages will, when composed with the voting strategies of the watchdogs, produce the watchdog vote vector; when composed with insider voting strategies, they produce the insider vote vector. This process of composition generates a map, from strategies to votes, which we represent as follows:

$$\mathbf{V}(\sigma^W, \sigma^I(s)) = \Big(\mathbf{v}^W\big((\mathbf{m}^W \mathbf{m}^I(s))\big), \mathbf{v}^I\big((\mathbf{m}^W \mathbf{m}^I(s)), s\big)\Big).$$

*2.4 Payoffs*

The payoffs to the agents depend on four events: (1) the signal received by insiders, $\tilde{s}$, (2) whether insider voting exhibited consensus, (3) whether the project was approved $(A)$ or rejected $(R)$, and (4) a variable $\tilde{z}$ equal to one if the penalty is assessed on insiders upon a consensus failure.

Voting exhibits consensus if no insider voted for an investment policy different from the policy adopted. Technically, consensus is defined in the model as follows. Let $c: V \to \{0, 1\}$ be the indicator function for whether all insiders support the majority consensus. That is, $c(v) = 1$, if and only if $\#\mathcal{Y}(v) > \#\mathcal{N}(v) \Rightarrow \#\mathcal{N}(v_I) = 0$ and $\#\mathcal{Y}(v) \leq \#\mathcal{N}(v) \Rightarrow \#\mathcal{Y}(v_I) = 0$.

The project is accepted if strictly more votes are cast for the project than against the project. We represent the acceptance with an indicator function $a$. This function, $a: V \to \{0, 1\}$, is defined as follows, $a(v) = 1$, if and only if $\#\mathcal{Y}(v) > \#\mathcal{N}(v)$.

The penalty for the failure of consensus is stochastic. Whether or not the penalty is imposed depends on a zero-one valued random variable $\tilde{z}$. This random variable is independent of $\tilde{s}$. The probability that $z = 1$ equals $\rho$. If consensus is not achieved, and $z = 1$, a penalty, $P > 0$, is imposed on all insiders.

Given these discussions, we can write the ex post payoff of agent $j$ which we represent by $U_j$ as follows: If $j$ is an insider then

$$\text{(DPI)} \qquad U_j(v,\omega) \equiv U_I(v,\omega) \equiv a(v)\,x(I,A,s(\omega)) + (1-a(v))\,x(I,R,s(\omega)) - P\,(1-c(v))\,\tilde{z}(\omega)$$

If $j$ is a watchdog then

$$\text{(DPW)} \qquad U_j(v,\omega) \equiv U_W(v,\omega) \equiv a(v)\,x(W,A,s(\omega)) + (1-a(v))\,x(W,R,s(\omega))$$

The rankings of the payoffs, $x$, given in the above equations are as follows

$$\text{(A.1)} \qquad \begin{gathered} x(I,A,G) > x(I,R,G),\ x(I,A,B) > x(I,R,B); \\ x(W,A,B) > x(W,R,B),\ x(W,A,B) < x(W,R,B). \end{gathered}$$

$$\text{(A.2)} \qquad \pi\,x(W,R,G) + (1-\pi)\,x(W,R,B) > \pi\,x(W,A,G) + (1-\pi)\,x(W,A,B)$$

$$\text{(A.3)} \qquad \rho\,x(I,A,s) - \rho\,P > x(I,R,s),\ \ s = G \text{ or } B$$

Assumption A.1 expresses the fact that insiders prefer acceptance of the project regardless of its quality and watchdogs prefer to accept the project only when it is good, i. e., $s = G$. A.2 implies that, if watchdogs have to make their accept/reject decision based on their prior information, they prefer to reject the project. A.3 implies that insiders are willing to pay the price for lack of consensus if, by paying this price, they will ensure acceptance of the project. An implementation policy that is consistent with the preferences of watchdogs is called an efficient policy. This terminology is natural given our interpretation that watchdogs represent the uninformed principals.

Using the definition of agent utility given in (DPI) and (DPW), we see that the utility of an agent $j$ under strategy vector $\sigma$ is given by

$$u_j(\sigma) = \mathrm{E}\big(U_j(\mathbf{V}(\sigma^W,\sigma^I),\omega)\big).$$

*2.5 Results*

First, we consider coalition-proof equilibria. As we show in the Appendix, coalition proofness ensures that insiders and watchdogs each act as if each group were a single agent. This restriction implies that insiders will force acceptance of the project whenever acceptance is not blocked by watchdogs. Knowing that insiders will not condition their support for the project on the information signal, watchdogs realize that they must choose between accepting the project for both information signals or rejecting the project for both information signals. By Assumption A.2, watchdogs prefer blocking the project for both signals to accepting the project for both signals. Thus, watchdogs will always block the project. Realizing this fact, insiders will also vote against the project, so as not to call down the penalty for lack of consensus. The next theorem summarized these results.

**Theorem 1.** *In all coalition-proof equilibria, (a) all insiders vote to reject the project under both information signals and (b) under both information signals, enough watchdogs vote against the project to ensure that even if all insiders were to switch their votes to acceptance, the project would still be rejected. Moreover, coalition-proof equilibria exist in which all agents vote against the project under all information signals.*

*Proof.* See Appendix.

In the coalition-proof equilibrium, watchdogs and insiders will not deviate to an equilibrium in which watchdogs abstain and the efficient project choice is implemented, because watchdogs realize that an agreement to switch to this equilibrium would be betrayed by insiders. Thus, watchdogs veto. Given the watchdog veto, insiders recognize supporting the project is futile and thus themselves vote against it.

The consequences of this theorem for our experimental design are clear. In our experiments, we assume the firm has a board consisting of 4 outsiders and 3 insiders. Theorem 1, specialized to this context, asserts that, in all coalition-proof equilibrium outcomes,

a. All three insiders vote to reject the project under both information signals;

b at least three outsiders vote to reject the project under both information signals;

c. the project is rejected with probability 1, and insider consensus is never violated.

Next we define efficient equilibria.

**Definition 1.** *An outcome is efficient if (a) the project is accepted only when information signal $G$ occurs and rejected only when signal $B$ occurs and (b) there is no chance that insiders will incur the penalty for a failure of consensus.*

We now show that Nash equilibria producing the efficient outcome exist. These, these equilibira are not coaltion-proof. However, Nash equilibrium requirement only requres that agents consider the effects of unilateral deviations from candidate equilibrium strategy vectors. Insiders know that if other insiders are providing "honest" recommendations by sending different messages depending on the information signal, unilateral efforts of a single insider to ensure project acceptance under both signals is futile and will simply call down the penalty for consensus failure. Thus, insiders will not deviate from the candidate equilibrium. Because the candidate equilibrium produces the highest possible payoff to outsiders, they will not deviate from the equilibria implementing the efficient outcome.

**Theorem 2.** *In all Nash equilibria that implement the efficient outcome all insiders vote $\mathcal{Y}$ if they observe the signal $G$ and vote $\mathcal{N}$ if they observe the signal $B$. Moreover a Nash equilibrium implementing the efficient outcome exists.*

*Proof.* First note that consensus among insiders is necessary in all Nash equilibria that implement the efficient outcome. This follows because a lack of unanimity among insiders would result in the possibility that insiders would bear a penalty. Thus, in all equilibria that implement the efficient outcome, all insiders vote $\mathcal{Y}$ if they observe the signal $G$, and vote $\mathcal{N}$ if they observe the signal $B$.

Next we show that an equilibrium exists that implements the efficient outcome. The equilibrium is given as follows. All agents send the same arbitrary message $m_o$, that is, $\mathbf{m}_i^W = \mathbf{m}_i^I(s) = m_0$. All insiders follow the strategy of voting $\mathcal{Y}$ if and only if they receive information signal $G$; $\mathbf{v}_i^I(m, G) = \mathcal{Y}$, $\mathbf{v}_i^I(m, B) = \mathcal{N}$. All watchdogs abstain regardless of the messages they observe, $\mathbf{v}_i^W(m) = \mathcal{A}$. In this candidate equilibrium, the vote of an individual agent cannot change the project selected. Moreover, for insiders, deviation from the strategy may incur the penalty for lack of consensus. Thus, unilateral deviations from the candidate information strategy vector cannot increase the payoff to any of the agents. It follows that the candidate strategy vector is a Nash equilibrium. □

In the efficient Nash equilibrium, the fact that all other insiders vote against the project when the information signal is $B$ ensures that no individual insider can change the project selected by changing her vote. In fact, a unilateral vote change by a single insider will not only have no impact on the outcome of the vote, it will also subject insiders to the penalty for lack of consensus. Theorem 2, when translated into the specific context of our experiments, shows that the outcomes of efficient equilibria all share the following characteristics.

a. All three insiders vote for the project under information signal $G$ and against the project under information signal $B$.

b. Watchdogs tend to be passive in that at least two watchdogs abstain when the information signal is $G$.

c. The project is accepted if and only if the information signal is $G$.

## 3 Experimental design

The subjects in the experiment were undergraduate finance majors and MBA students who were currently enrolled in introductory corporate finance classes. They were told they would have an opportunity to earn money in a research experiment involving group decision-making. The entire experiment consisted of several treatments. For each treatment we conducted between one and three experimental sessions each lasting ten rounds. Eevry subject participated in only one experimental session.

### *3.1 The basic design*

We now describe the basic features that were common to all treatments. Variations of this basic design are discussed below. At the beginning of each session, the subjects were read a set of instructions (see Appendix A); they completed relevant worksheets, and were given the opportunity to ask questions. At the end of this instructional period, the monitors randomly divided the subjects into groups of seven.

Next, subjects were randomly assigned their agent type—insider or watchdog.[8] With the exception of one treatment, each group consisted of three insiders and four watchdogs. In the one exception, all subjects were classified as insiders. The group size and the number of insiders was the minimum number needed to ensure that (i) the defection of one insider from a unanimous vote by insiders did not cancel the insiders' majority and (ii) outsiders had enough votes to override insiders.[9]

Groups then dispersed to different ends of a large classroom. Each round began with the group-subgroup communication protocol: First discussion was permitted among all seven members of each group; then, after each group split into subgroups by agent type, another discussion followed. The following restrictions applied: No physical threats, no side payments, no talking between groups, and a maximum of four minutes for each group discussion. Subjects never appeared to find this time limit to be binding.

Next, the insiders from each group watched a monitor draw the project outcome from a bucket. To ensure that good and bad draws had equal probabilities, the bucket contained 50 white chips (good outcome) and 50 red chips (bad outcome). Chips were replaced after each draw. Following a draw, the insiders returned to their groups. Discussion among all members of each group was permitted once the "informed" insiders had returned to their respective groups. The time limit for this discussion was two minutes but was never appeared to be binding. After the discussion, all subjects in each group privately cast their ballots.

---

[8] The terms "insiders", "watchdogs", and "project" were never mentioned in the instructions or used orally during the experiment. Instead, the insiders were always referred to as "Type A" participants, the watchdogs as "Type B" participants, and the project as the "event." The term "participants" was used instead of "agents."

[9] This is in keeping with Jensen (1993), among other researchers, who has suggested that board size should be limited to seven or eight members, so that the marginal cost of coordination and processing problems does not exceed the marginal benefit of the additional members' input.

Insiders could cast either a yes or a no vote, while watchdogs could either abstain or vote no. After all the subjects had voted, a monitor counted the votes. The project was assumed to have been undertaken if the yes votes outnumbered the no votes. Split votes, when the number of yes votes equaled the number of no votes, were deemed to result in the rejection of the project. The monitor privately informed each group of the aggregate vote for the group, and of the distribution of votes by agent types after tallying the votes and ascertaining the outcome. The outcome, together with the project type and the occurrence of a split vote, determined payoffs for the period. Most treatments incorporated a penalty feature for split votes. If at least one insider's vote disagreed with the votes of other insiders or did not conform with the majority vote for the group, a monitor drew a chip from a bucket of poker chips that contained 20 blue chips and 80 white ones. Chips were replace! d after each draw. Insiders in the group were assessed a penalty if a blue chip was drawn.

Each round ended after subjects learned about the outcome, participated in the penalty draw if applicable, and calculated their earnings. Each experimental session consisted of ten periods, but subjects essentially played a game with an indefinite endpoint since they were not told how many periods they would play.

Payoffs were designed to ensure that insiders preferred to take on the project regardless of the outcome (Assumption A.1). In the absence of a penalty, they received at least \$0.90 following a majority yes vote, compared with a maximum \$0.60 following a majority no vote. To ensure that the penalty mechanism was material, we included a 20% probability that a penalty would be imposed on all insiders if their votes lacked consensus. When a penalty was imposed, insider payoffs fell by \$0.35. Thus, insiders could expect a payoff of at least \$0.65 following a majority yes vote even after a penalty was imposed. Because this amount was higher than their expected \$0.60 payoff from a majority no vote even in the absence of a penalty, the penalty was not sufficient to reverse their preferences between investing in the project and rejecting it (Assumption A.3).

Watchdogs payoffs were designed to ensure that they preferred taking on the project only when it was good. More specifically, the watchdogs could expect to earn \$0.70 from investment in the project conditional on a good draw, and \$0.00 from investment conditional on a bad draw.

Thus, they expected a payoff of \$0.35 for a majority yes vote, which is less than the expected payoff of \$0.50 for a majority no vote (Assumption A.2). The payoff structure for both agent types was common knowledge.

*3.2 The central treatments*

We now describe the five central treatments. These treatments were designed to examine the importance of two central features of the mechanism described above—watchdogs and penalties for split votes. We varied mixing protocols because theoretical and experimental research has shown they can influence experimental outcomes.[10]

Table 1 summarizes the salient characteristics of the thirteen sessions employing the five central treatments.[11]

As the first three rows of the table indicate, an important shared feature Among these treatments was the ability of agents to talk to each other before each vote. As we explained in the previous section, prior to drawing for project quality, communication was first permitted among all members of a group, and then within sub-groups based on agent type. After the drawing and before voting, "informed" insiders were allowed to communicate with "uninformed" watchdogs in their group.

---

[10] For example, some mixing protocols could foster learning about other subjects' beliefs about the game and enhance cooperation. However, Issac, McCue, and Plott (1985), demonstrate that repetition decreases cooperation in public goods experiments, and Eckel and Holt (1989) find that strategic voting in an agenda-setting context is more likely to occur in committees that interact over time on related issues.

[11] We planned to run 3 sessions in each of the no-penalty experiments. However, even though we over-recruited by five subjects in the first non-penalty experiment, not enough students showed up to form the third group of seven. Given our budget constraint and the fact we would have at least 20 observations with the two sessions, we decided to go ahead and run the experiment.

Row 4 of Table 1 demonstrates that the five central treatments fall into two groups based on the mixing protocols employed. In the first group, which consists of three treatments over eight sessions, we employed the random mixing protocol. Here, after voting in each period, subjects were randomly mixed into a new group but maintained their agent type for the entire session. This mixing protocol most accurately captures the essence of the theoretical model developed above, in which the vote is modeled as a single shot game. In real world situations however, membership of boards remains stable across a number of votes. To examine the relevance of our model to such situations, we also employed a repeated groups protocol which fixed group membership for an entire session. Two of the centraal treatments, consisting of a total of five sessions, employed this protocol.

Our results stress the usefulness of multi-agent mechanisms, whose merits we attempted to directly gauge. On of our random mixing protocol treatments (the exception noted earlier) only allowed for insiders, i.e., 7 insiders and 0 watchdogs in each of three sessions. The other five sessions using the random mixing protocol employed the standard set-up described earlier, i.e., 3 insiders and 4 watchdogs in each session.

The penalty for split votes plays an important role in the theory outlined above because it makes defections from the consensus vote costly for insiders. Thus, even single-agent mechanisms can implement the efficient policy. To examine the impact of the penalty, for each of the two mixing protocols, we ran two sessions where the insiders did not face a penalty draw after a split vote.

*3.3 Robustness*

Numerous experimental studies have demonstrated that communication and "cheap talk" in increase efficiency, though other studies suggest that the rule structure of permitted communication and the complexity of the social dilemma setting can affect the robustness of communication (see, e.g., Farrell and Rabin, 1996).[12] Because theoretical work also suggests that cheap-talk pro-

[12] As Isaac and Walker (1988) suggest, the role of communication may be (1) to speed members' awareness of optimal group behavior—implying that the effects of communication will continue even if the ability to communicate is later removed; or (2) to influence members' beliefs about

tocols matter (see, e.g., Milgrom and Roberts (1996)), we considered different discussion-sequence protocols in our robustness sessions. We ran nine treatments, distributed over eleven sessions, to examine the robustness of the results from the central treatments. The robustness treatments are described below and a summary of their features is presented in Table 2.

The first six treatments were identical to the central treatment employing the repeated groups mixing protocol with penalties for split votes in all but the communication protocols. In the first robustness treatment, applied over three sessions, subjects were not permitted to communicate after they had been assigned to their groups. Only one session was run for each of the remaining robustness treatments.[13] The second robustness treatment employed the subgroup-group protocol for communication prior to the draw for project quality—subgroups of insiders and watchdogs first discussed strategies separately before joining their groups to continue the discussions. This reversed the sequence of communication in the central treatments.

Each of the next four robustness treatments varied the sequence of communication prior to the draw for project quality and did not allow for verbal communication between subgroups following the draw for project quality. In the third robustness treatment we allowed for an additional period of communication between agent sub-groups before the project-quality draw—the subgroup-group-subgroup sequence of communication. The fourth robustness treatment employed the subgroup-group sequence communication protocol prior to the project quality draw, just as did the second robustness treatment described above. The fifth treatment employed the group-sub-group sequence of communication employed in the central treatments. The sixth only allowed for communication within the group as a whole, but prohibited subgroup communication.

---

other members' ongoing responses—implying that communication must continue if it is to be effective. This distinction may be important to governance organizations where the cost of getting participants together for all decisions is prohibitive.

[13] Because these treatments were merely exploratory in nature, we ran only one session each to investigate the influences of changes in protocols on outcomes.

In the latter four robustness treatments, for the first five periods, communication between "informed" insiders and "uninformed" watchdogs following the draw for project quality took the form of the insider subgroup passing a piece of paper to the watchdog subgroup via a monitor.[14] On the piece of paper, the insider subgroup was permitted to convey its strategy only by circling either yes or no. During the final five periods, no communication was permitted following the project quality draw.

Robustness treatments seven through nine, the final three robustness treatments, mirrored the robustness treatments four, five, and six described above with one exception—they employed a hybrid of the mixing protocols in the central treatments. After each round watchdogs were randomly shuffled across groups while insider subgroups remained unchanged. This randomized-watchdog mixing protocol examined the performance of the mechanisms described above when insiders are expected to have long tenures but watchdog membership is temporary. We designed the treatment to simulate insider board membership that remains relatively stable over long periods.[15]

## 4 Results from central treatments

In this section we present results from the central treatments. We examine these results along four dimensions—(1) the incidence of institutionally preferred outcomes, (2) insider voting patterns, (3) watchdog voting patterns, and (4) the congruence of voting vectors to those supporting the two competing equilibria. The predominance of efficient outcomes indicates that equilibria supporting efficient outcomes enjoy greater predictive success, and that watchdogs have the desired effect of greatly improving efficiency. While the penalty for split votes does not appear to have a perceptible

[14] A written message, which is not as easily misinterpreted (or strategically miscommunicated) as possibly face-to-face communication, should provide an environment more conducive to truth-telling and reputation building. For this reason, we analyzed the communication-mode protocol of a written message followed by rounds with no communication.

[15] Such a situation could be expected to prevail on the board of a corporation whose management is entrenched (see, e.g., Shleifer and Vishny (1989)).

effect on insider voting, switching to the repeated mixing protocol appears to encourage insiders to behave opportunistically. Occassionally, this behavior leads to voting patterns consistent with coalition-proof equilibria.

*4.1 Incidence of efficient policies*

Table 3 and Figure 1 present the frequency with which majority votes resulted in the adoption of institutionally preferred policies. Overinvestment typifies the outcomes of the treatment with no watchdogs. The project was rejected once in 30 draws and was never rejected following a bad draw.

Watchdogs perceptibly increased efficiency. This supports the belief that multi-agent mechanisms facilitate full implementation of efficient outcomes. In both treatments employing the random mixing protocol, majority votes always resulted in the institutionally preferred outcome. The adoption of the institutionally preferred policy fell off in the repeated group treatments. The decline was greatest following good draws. In the repeated group treatment with penalties for a split vote, only 60% of the votes following good draws resulted in the institutionally preferred outcome compared with 95% following bad draws. As Figure 2 demonstrates, this difference can be attributed primarily to one session of the repeated treatment with penalties in which, during the second half of the session, the project was rejected three times after good draws. The drop in the aggregate percentage of institutionally preferred outcomes following good draws, together with the relatively high p! ercentage of institutionally preferred outcomes following bad draws, suggests that the coalition-proof equilibria may also have some predictive power. However, given that the percentage of institutionally preferred outcomes in two of three sessions is comparable to that in the random mixing sessions, we can argue that equilibria supporting the efficient outcome continue to describe outcomes well.

We employ Chi-square tests to examine the impact of introducing watchdogs, changing the mixing protocol, and eliminating the penalty for split votes. The results, presented in Table 4, confirm that the presence of watchdogs significantly influenced the frequency of institutionally preferred outcomes following bad draws, and, in turn, suggest that multi-agent mechanisms can significantly increase the likelihood that institutionally preferred policies will be adopted. Contrary to expectations, however, a penalty for split votes has little effect. In both the repeated groups and the random mixing protocols, there is no evidence that the elimination of the penalty affected the percentage of institutionally preferred outcomes. The tests do, however, indicate significant differences in the outcome distributions of the repeated groups and random mixing treatments conditioned on good draws. The tests are significant at the 1% confidence levels for treatments employing penalties f! or split votes, and at the 10% level in the absence of penalties. This suggests that repeated interaction between agents may encourage the incidence of institutionally preferred policies.

*4.2 Insider votes*

Figure 3 and Table 5 present the frequency distribution of insider votes in the central treatments. A cursory examination reveals overwhelming insider support for the project in the treatment with no watchdogs. Deviation from this pattern occurred only twice in 18 draws, providing strong support for the predictions that single-agent mechanisms can result in the misallocation of resources.[16]

The move to a multi-agent mechanism appears to significantly alter insider voting patterns. In the two random mixing treatments with watchdogs, insider behavior conformed perfectly with voting strategies in equilibria supporting the efficient outcome. Insiders always voted unanimously to accept the project following a good draw and to reject it following a bad one. The penalty for a lack of consensus did not alter insider voting patterns.

Behavior in the repeated group treatments fell between the two extremes. The vast majority

[16] The first vote occurred after six rounds, when one subject convinced his group to vote against the project despite a good draw. This resulted in a unanimous vote to reject the project. A monitor overheard him suggest that there might be an additional reward for doing so, even though a monitor had assured the group otherwise during the instruction period. In the next period the same individual voted no again, despite a good draw. The remainder of the group, however, voted to accept.

of insiders displayed some form of opportunistic behavior. At least one insider voted to reject the project 30% of the time following good draws, and at least one insider voted to accept 28% of the time following bad draws. In four of the five sessions, insiders attempted to fool the watchdogs into accepting a project at the first bad draw. At least once in each of these four sessions, insiders followed the strategy of voting to reject following a good draw and to accept following a bad one. In aggregate, however, insider opportunism was evident in only 24% of the rounds. This is because only in two sessions did insiders attempt to fool the watchdogs on multiple occasions. In the remaining three sessions, following the first two draws, insider behavior never deviated from voting strategies that are consistent with equilibria supporting the efficient outcome. In contrast, as illustrated by Panel A of Figure 4, in the final periods of one session, unanimous negative votes by insiders followed good draws on two occasions. This provides additional evidence to support coalition-proof equilibria.

Chi-square test statistics to examine the effect of changes in the mixing protocol and the presence of penalties on insider voting are presented in Table 6. They support the hypothesis that the mixing protocol can influence voting. In the repeated group treatments without penalties, voting differed significantly from that in both random mixing treatments following bad draws. Following good draws, the mixing protocols do not appear to significantly influence voting. As with the outcome distributions, we find no evidence that the penalty for split votes significantly altered the distribution of votes. We arrived at a similar conclusion when we tested for differences in the distributions of unanimous insider votes across treatments with and without penalties.

### *4.3 Watchdog votes*

We now turn to an examination of watchdog voting. Figure 5 and Table 7 show that watchdogs usually voted unanimously to reject the project following bad draws and to abstain following good draws. These voting patterns suggest that insiders accurately transmitted their information about project quality to watchdogs. Consequently, watchdog votes tended to conform with strategies employed in equilibria supporting efficient outcomes.

Once again the mixing protocol appears to have influenced voting patterns. Watchdog votes in treatments employing the random mixing protocol provide strongly support the hypothesis that multi-agent mechanisms fully implement the efficient outcome. In treatments employing the repeated groups protocol, watchdog votes also supported the hypothesis that the coalition-proof equilibria have some explanatory power.

Voting patterns varied following good draws. In the two random mixing treatments, watchdogs coordinated their votes with the draws and always voted unanimously to abstain. In the random group treatments, watchdogs as a group were passive in that at least two watchdogs abstained 70% of the time when insiders faced the possibility of a penalty for split votes, and 88% of the time when there was no threat of a penalty. This is consistent with the voting patterns supporting efficient equilibria. In the repeated group treatments, however, watchdogs occasionally voted unanimously to reject the project following good draws, providing some support for the coalition-proof equilibria. When there were no penalties for split votes, unanimous rejection occurred once in 8 draws. In the treatments penalizing split votes, unanimous rejection occurred twice in 10 draws. As can be seen from Figure 6, these votes followed soon after votes in which insiders managed to secure watchdog cooperation in getting the project accepted following a bad draw.

Watchdog votes following bad draws (Panel B of Figure 5) suggest that watchdogs were more passive in the repeated treatments. They unanimously abstained from voting in 6 out of 32 (19%) draws. This never occurred in the random mixing treatments.

Chi-square tests indicate that changes in the mixing protocol significantly influenced watchdog behavior. The evidence presented in Table 8 supports the hypothesis that, following good draws, the change in mixing protocols influenced behavior in the treatments with penalties. The tests do not support the hypothesis that the penalty for split votes influenced watchdog behavior.

*4.4 Subject votes and equilibrium vote vectors*

Having described the outcomes and voting behaviors of subgroups, we now examine the congruence between subject votes and voting patterns that support the two competing sets of equilibria.[17]

[17] Note that for treatments employing only insiders, the coalition-proof equilibrium requires that all insiders vote unanimously to accept the project regardless of the draw. This follows because the coalition of all insiders maximizes payoff by accepting the project regardless of the signal. In

Figure 7 graphs the frequency with which the vote vector in the experiments conformed with those supporting the two competing sets of equilibria. This information also appears in Table 9.[18] Figure 7 also presents the frequency with which vote vectors do not conform with equilibrium vote vectors.

Subject votes more frequently resembled those supporting efficient equilibria. In fact, they conformed with those supporting efficient equilibria over 60% of the time. The two random grouping treatments including watchdogs enjoyed the highest possible success rate. Confirming the earlier evidence on the impact of the mixing protocol, voting behavior in the repeated group treatments bore a weaker resemblance to voting strategies supporting efficient equilibria. Further, the coalition-proof equilibrium enjoyed its highest success rate in these treatments.

The success rate of the two competing equilibria conditioned on draws is presented in Figure 8. Panel A illustrates success following good draws, while Panel B does the same for bad draws. From this figure it is clear that the success enjoyed by the coalition-proof equilibria tends to be concentrated in votes following bad draws, but appears to have no predictive power following good draws in the two random mixing treatments with watchdogs. A simple explanation for this

---

the efficient equilibria it is still the case the insiders unanimously vote to accept following a good draw and reject following a bad one.

[18] The definitions employed to classify vote vectors for treatments with watchdogs are based on Theorems 1 and 2 above. These definitions also classify vote vectors in treatments where splits votes do not result in the possibility of a penalty. However, in these cases, additional vote vectors may support efficient and coalition-proof equilibria. For example, in the absence of the threat of a penalty, certain coalition-proof equilibria allow for at least three watchdogs to vote to reject the project regardless of the draw while insiders unanimously vote to accept. Eliminating the penalty removes the requirement for consensus. The use of a broader definition only changes the results presented for the repeated treatment with no penalty following a bad draw. In this case, the proportion of vote vectors satisfying the requirements of those supporting efficient (coalition-proof) equilibria rise from 7/12 (7/12) to 11/12 (11/12).

phenomenon is that, following bad draws, some of the efficient equilibria are supported by voting formations that are identical to those supporting the coalition-proof equilibrium. Thus, following bad draws it is not always possible to allocate subject strategies to one of the two competing equilibria. However, the appreciable difference in the success enjoyed by the two equilibria following good draws supports the view that subjects tended to play strategies supporting efficient outcomes.

## 5 Robustness

In this section we present evidence from our robustness sessions. We focus on the influence of changes in mixing and communcation protocols on the efficiency of outcomes and the congruence between subject votes and equilibrium vote vectors. The evidence indicates that our results are relatively insensitive to these changes.

The frequency with which the efficient policy is adopted in the robustness treatments suggests that changes in communication protocols had little effect. Table 10 indicates that, with the exception of the treatment with no subject communication, the percentage of votes ressulting in the adoption of th efficient policy in the robustness treatments uniformly exceeds that obtained in the central treatment employing the repeated groups with penalties protocol. The conditional distributions of outcomes also produced this result. Statistical tests for differences in the outcome distributions of the robustness treatments and the central treatment employing the repeated groups with penalties protocol are presented in Table 11. This table provides little support for the hypothesis that the outcome distributions are sensitive to changes in treatments.

Table 12 presents the frequency with which subject vote vectors satisfied requirements of equilibrium vote vectors. Once again the efficient equilibrium appears to account for subject behavior. With the exception of the robustness treatment with no subject communication, at least 80% of subject votes resembled those that support efficient equilibria. Subject votes resembled those supporting coalition-proof equilibria about 20% to 30% of the time. In most cases these votes followed bad draws. Again, then, the apparent power of the coalition-proof equilibria to explain subject votes following bad draws could merely be driven by the same vote vectors that also support a number of efficient equilibria. However, as with the core treatments, there was some evidence that coalition-proof equilibria could explain votes following good draws.

In the robustness treatment without subject communication, the equilibria displayed diminished explanatory power. Vote vectors resembled those supporting efficient equilibria 30% of the time, but subject behavior could not be explained by the coalition-proof equilibria. This supports the idea that in the absence of communication subjects more likely will gravitate to equilibria that require less coordination—the efficient equilibrium.

## 6 Conclusion

We implement a voting mechanism in a setting where a group of insiders, whose incentives are not necessarily aligned with those of the institution, have private information regarding the institutionally preferred allocation. Our laboratory analysis finds that the presence of uninformed "watchdog" agents, whose votes outnumber the insiders', can reduce the incidence of undesirable equilibria and improve organizational performance. Inefficient equilibria, however, cannot be completely eliminated. Contrary to expectations, institutionally preferred allocations are more likely to arise when group membership is frequently altered. In addition, indirect evidence suggests that a board size of seven, as suggested by Jensen (1993) among others, is sufficient for efficient decision making.

Gilson and Kraakman (1991) suggest that independent directors, in order to be effective, should not be merely independent of management but accountable to shareholders. Empirical studies use proxies for the alignment of incentive structures of independent directors with the institution, which are by their nature noisy. For example, though some directors may be classified as independent of management, they may be beholden to management in subtle ways, such as by acting as paid advisors or consultants to the company. Our experiments contribute to the empirical research in this area by providing A framework that can control for conflicts of interest and ensure that the independent directors are, truly, watchdogs.

The role of watchdog agents in many organizational structures and in our study has been primarily to vote on designated issues, but the positive role watchdogs play in attaining the institutionally preferred allocation in this governance structure may extend to a broader agenda-setting context. Other experimental studies have examined specific agenda-setting issues. An interesting area of further study would be to extend the research in both these areas to the role of independent directors within various agenda-setting environments.

## 7 Appendix

This appendix provides the formal proofs of Theorems 1. Before we initiate our proofs we require some definitions.

### 7.1 Definitions

**Definition 2.** *A strategy $\sigma'_j$ is a* best response *to $\sigma' \in \Sigma$ if the following implication holds for all $\sigma \in \Sigma$*

$$\forall k \in [N] - \{j\},\ \sigma_k = \sigma'_k \Rightarrow u_j(\sigma') \geq u_j(\sigma).$$

**Definition 3.** *A strategy vector $\sigma^*$ is a* Nash equilibrium *if for all $j \in [N]$, $\sigma^*_j$ is a best response for $j$ to $\sigma^*$*

**Definition 4.** *A* coalition-proof Nash equilibrium *is defined by induction on the size of coalitions as follows. Let $S \subset [N]$*

*(i) Suppose #(S) = 1, then $S = \{j\}$ for some $j \in [N]$. In this case, $\sigma$ is optimal for $S = \{j\}$ if and only if $\sigma_j$ is a best response to $\sigma$ for $j$.*

*(ii) Assume optimality has been defined for all $S$ such that $\#(S) \leq k-1$. Define optimality of coalitions $S$ of size $k$ as follows:*

*(a) $\sigma$ is self-enforcing for $S$ if $\sigma$ is optimal for $T$, whenever $T$ is a strict subset of $S$.*

*(b) $\sigma$ is optimal for $S$ if it is self-enforcing for $S$ and there does not exist any strategy vector $\sigma'$ which is also self-enforcing for $S$ such that*

$$\forall j \in [N] - S,\ \sigma'_j = \sigma_j, \tag{2}$$

$$\forall j \in S,\ u_j(\sigma') > u_j(\sigma). \tag{3}$$

*Finally, if $\sigma$ is optimal for $[N]$, we say that $\sigma$ is a* coalition-proof Nash equilibrium.

**Definition 5.** *If $\sigma$ is a strategy vector and $K$ is a subset of agents containing at least one insider, then we say that the insiders in $K$ are* decisive for $\sigma$ for signal $s$ *if, holding the actions of all other* *agents fixed, they can force acceptance of the project. In other words,*

$$\#(K \cap [I]) + \#\{j \in [I] - K \colon \mathbf{V}_j(\sigma^W, \sigma^I(s)) = \mathcal{Y}\} >$$
$$\#\{j \in [W] \colon \mathbf{V}_j(\sigma^W, \sigma^I(s)) = \mathcal{N}\} + \#\{j \in [I] - K \colon \mathbf{V}_j(\sigma^W, \sigma^I(s)) = \mathcal{N}\}.$$

Insiders belonging to a subset of agents are decisive for a given information signal if, by collectively changing their vote to yes when they receive that signal, they can ensure that the project is accepted. Next note that all insiders have the same payoff function, and that all watchdogs have the same payoff function. Thus, the strategy vector that maximizes the payoff to a subset of agents that consists only of insiders or outsiders is well defined. This motivates the following definition.

**Definition 6.** *Let $\sigma$ be a strategy vector, let $K$ be a nonempty "pure" subset of agents consisting only of insider types or only of outsider types. Suppose that over all strategy vectors $\sigma'$ such that $\sigma'_j = \sigma_j$, $j \notin K$, $\sigma$ produces the highest payoff to agents in $K$, then we say that $\sigma$* maximizes payoffs over *$K$.*

*7.2 Lemma used to establish Theorems 1 and 2*

Our most important lemma, Lemma 1, is quite straightforward. It implies that in coalition-proof outcomes pure coalitions consisting of just insiders or just outsiders act as if they are a single agent, maximizing their collective payoff over their joint strategy space.

**Lemma 1.** Let $\sigma$ be a strategy vector and let $K$ be a nonempty "pure" subset of agents consisting only of insiders types or only of outsider types. The strategy $\sigma$ is optimal for $K$ if and only if $\sigma$ maximizes payoffs over $K$.

*Proof.* Our proof is based on induction on the size of the coalition. If the coalition size, which we represent by $k$ equals 1, the lemma follows from the definition of a Nash equilibrium.

Next, suppose that Lemma 1 holds for a subset of size less than or equal to $k$. Consider a pure subset $K$ of size $k+1$ and a strategy vector, $\sigma$, that maximizes type payoffs over $K$. All subsets of a pure subset must be pure. Maximizing a type's payoff over a subset of $K$ can never yield a higher payoff than the payoff from maximizing over $K$. Thus, $\sigma$ must be self-enforcing for

$k+1$. Because $\sigma$ maximizes over $K$, no other strategy vector produces a higher payoff. Thus, $\sigma$ is optimal for $K$.

To prove the other leg of the if-and-only-if assertion, suppose that $K$ is pure coalition of size $k+1$ and let $\sigma$ be a strategy vector that does not maximize type payoffs over $K$. Then there must exist a strategy vector, say $\sigma' \neq \sigma$, such that $\sigma'$ equals $\sigma$ for agents not in $K$ and $\sigma'$ maximizes the payoff over $K$. (Because there are only a finite number of distinct strategy vectors, existence of a maximizing vector is guaranteed.) By the results of the previous paragraph, $\sigma'$ is optimal, and thus, a fortiori, self-enforcing for $K$. This implies that $\sigma$ cannot be optimal for $K$. Thus, maximization of the type payoffs over strategies in $K$ is necessary condition for optimality as well.
□

By Lemma 1, insiders will force project acceptance except when project acceptance is blocked by insufficient watchdog votes. Because information regarding the information signal is transmitted only by informed insiders, insiders can always force acceptance under one signal if they can force acceptance under any signal. This reasoning underlies the next lemma, Lemma 2.

**Lemma 2.** *If $\sigma$ is any coalition-proof Nash equilibrium, $\sigma$, the project is accepted with probability 1 or probability 0.*

*Proof.* Suppose that, under $\sigma$, the project is accepted under signal $s_1$ but not under signal $s_2$. Consider the subset consisting of all insiders. Consider the strategy vector $\sigma'$ calling for insiders to use the message and voting strategy that they use when receving $s_1$ under $\sigma$ for both information signals. Under $\sigma'$ the project is accepted with probablity 1. Assumptions A.1-A.3 ensure that this outcome produces a higher insider payoff than strategy $\sigma$. Thus, $\sigma$ cannot maximize the payoffs over $[I]$. Thus, $\sigma$ is not optimal for $[I]$ and thus is not coalition-proof. □

Because the payoff to watchdogs is always higher if the project is rejected under both signals than it is if the project is accepted under both signals, the coalition of all watchdogs can always gain by forcing universal rejection in any candidate equilibrium in which the project is being accepted with probability 1. Thus, such equilibria are not coalition-proof.

**Lemma 3.** *In any coalition-proof equilibrium, insiders are not decisive under either information signal, i.e., under both signals watchdogs cast sufficient votes against the project to block passage regardless of the votes of the insiders.*

*Proof.* Consider a coaltion-proof Nash equilibrium $\sigma$. By Lemma 2 we know that if the project is accepted at all, then the project is accepted with probability 1. Thus, the equilibrium payoff to watchdogs is $\pi\, x(W, A, G) + (1 - \pi)\, x(W, A, B)$.

Now suppose watchdogs deviate to the strategy of always voting against the the project, that is, consider the strategy vector $\sigma'$ defined as follows. For insiders, play the strategies prescribed by $\sigma$; for outsiders, play the message strategies prescribed by $\sigma$ but follow the voting strategy of voting against the project regardless of the message sent in the message phase. Because watchdogs outnumber insiders, the project is rejected with probability 1. Thus, the strategy vector $\sigma'$ yields watchdogs a payoff of $\pi\, x(W, R, G) + (1 - \pi)\, x(W, R, B)$. By A.2, this exceeds the equilibrium payoff under $\sigma$. Thus, $\sigma$ does not maximize watchdog payoffs over $[W]$. By Lemma 1, $\sigma$ is not optimal for $[W]$ and thus $\sigma$ is not coalition-proof. □

Because the project is being rejected in all coalition-proof equilibria regardless of how insiders vote, and because of the penalty imposed on insiders when consensus fails, insiders collectively have an incentive to vote unanimously against project acceptance when their votes are not decisive. Thus, coalition-proof outcomes are characterized by unanimous insider rejection.

**Lemma 4.** *In any coalition-proof equilibrium, all insiders vote to reject the project.*

*Proof.* To obtain a contradiction, let $\sigma$ be a coalition-proof equilibrium in which not all insiders vote to reject the project. From Lemma 3 we know that in any coalition-proof equilibrium the project is voted down regardless of the insiders' voting behavior. Consider the subset of agents consisting of all insiders. If these insiders deviate to a strategy of voting against acceptance regardless of the messages sent in the message phase, the deviant strategy, by eliminating the possibility of the penalty for a lack of consensus, produces a higher payoff than the strategies insiders are playing under $\sigma$. Hence, $\sigma$ does not maximize payoffs for $[I]$. Thus, by Lemma 1, $\sigma'$ is not optimal for $[I]$ and hence $\sigma$ is not coalition proof. □

Lemmas 1 to 4 characterize coalition-proof equilibria. We now turn our attention to proving that coalition-proof equilibria exist. The existence proof requires us to consider mixed insider–watchdog coalitions. Characterizing such coalitions motivates the following definitions

**Definition 7.** *A coalition of agents, $K$, is* flawed *under strategy vector $\sigma$ if, there exists an information signal $s$ such that the following conditions hold.*

a. *The project is be rejected under $s$; that is, for some $s$,*

$$\#\mathcal{Y}(\mathbf{V}(\sigma^W, \sigma^I(s))) \leq \#\mathcal{N}(\mathbf{V}(\sigma^W, \sigma^I(s))).$$

b. *The coalition $K$ contains at least one insider.*

c. *The insiders in $K$ are decisive for $\sigma$ under $s$.*

A flawed coalition contains a subset of insider agents who are decisive for project acceptance yet fail to ensure that the project is always accepted. In the subsequent analysis we will show that equilibria in which the set of all agents is flawed are not coalition-proof. Because coalition proofness is defined by induction on subset size, we must define flawed coalitions not only for the set of all agents, but also for all proper subsets of agents. The next lemma shows that, when the strategy vector is flawed, by collective changes in their strategy, a sufficiently large coalition of insiders can always modify their strategies to ensure project acceptance.

**Lemma 5.** *If $J$ is flawed for $\sigma$ and if all insiders not in $J$ are sending the same message under both signals, there exists a strategy vector, $\sigma'$, which specifies the same strategies as $\sigma$ for all watchdogs and insiders not in $J$ such that the project is accepted with probability 1.*

*Proof.* We construct the strategy vector as follows. Let $K = [I] \cup J$, for all agents not in $K$; let $\sigma' = \sigma$. Next, determine the signal, say $s'$, under which insiders are decisive. Each agent in $K$ should (a) send under both signals the message that, under $\sigma$, she sent under $s'$ and (b) subsequently vote to accept the project regardless of the pattern of messages received. This strategy will ensure that the project is accepted under both signals with the unanimous support of insiders in $K$. $\square$

**Lemma 6.** *If $J$ is flawed for $\sigma$ and if all insiders not in $J$ are sending the same message under* *both signals, then $\sigma$ is not optimal for $J$.*

*Proof.* By Lemma 5, $\sigma$ does not maximize the payoff over $[I] \cup J$. Thus, by Lemma 1, $\sigma$ is not optimal for $J$. □

**Lemma 7.** *There exists a coalition-proof Nash equilibrium under which the project is rejected with probability 1 and all insiders cast votes against project acceptance under both information signals.*

*Proof.* Consider the strategy vector $\sigma$ defined as follows. All agents sent the same arbitrary message, $m_o$, independent of the information signal, i.e., $\mathbf{m}_i^W = \mathbf{m}_i^I(s) = m_0$. All insiders and watchdogs follow the strategy of voting $\mathcal{N}$ regardless of the information signal, that is, $\mathbf{v}_i^W(m) = \mathcal{N}$, $\mathbf{v}_i^I(m, G) = \mathcal{N}$, $\mathbf{v}_i^I(m, B) = \mathcal{N}$. To show that this is a coalition-proof Nash equilibrium, we need to show that $\sigma$ is optimal for all subsets of $[N]$. To show this we provide a proof by induction on subset size. First note that when the subsets contain a single element, optimality simply requires that for all agents $j$, $\sigma_j$ is a best response to $\sigma$ for $j$. However, no individual agent can change the project acceptance decision by unilaterally changing his strategy. Moreover, insiders may call down a penalty if they deviate from the consensus. Thus, the assertion of optimality holds for all subsets of size 1. Next suppose that, for subsets of size less than or equal to $k$, optimality holds. Consider a subset $K$ of size $k+1$. Given the induction hypotheses and the definition of coalition proofness, optimality requires that there not exist another self-enforcing strategy for $K$ that yields all the agents in $K$ a higher payoff. If $K$ consists only of watchdogs, $K$ cannot produce a higher payoff because, given the uninformative messages of insiders, watchdogs cannot induce a strategy vector that accepts the project under the good information signal and rejects the project under the bad signal. Given assumptions A.1-A.3, rejecting the project under both signal produces a higher payoff to watchdogs than accepting the project under both signals. No improving vector of strategies exists for $K$, and thus, a fortiori, no self-enforcing vector of strategies exists for $K$ when $K$ consists of a set of watchdogs. Next note that a coalition of all insiders does not have sufficient votes to change the outcome. Thus, such a subset cannot increase its welfare by deviating from the equilibrium. Only a mixed coalition, by implementing a vector of strategies calling for rejection when the information signal is $B$ and acceptance when the information signal is $G$, can increase the payoff to all agents in $K$. However, by our earlier definition, such a coalition is flawed. Lemma 6 shows that a flawed coalition is not optimal and thus, a fortiori, is not self-enforcing. Thus, optimality for coalitions of size $k+1$ has been established, proving the assertion of the theorem by induction. □

**Proof of Theorem 1.** This theorem follows directly from Lemmas 4, 5, 6, and 7. □

REFERENCES

Agrawal, A., and C. Knoeber, 1996, Firm performance and mechanisms to control agency problems between managers and shareholders, *Journal of Financial and Quantitative Analysis* 3, 377-397.

Bernheim, D., B. Peleg, and M. Whinston, 1987, Coalition-proof equilibria I: Concepts, *Journal of Economic Theory* 42, 1-12.

Bhagat, S., and B. Black, 1997, The uncertain relationship between board composition and firm performance, working paper, Columbia Law School.

Brickley, J., J. Coles, and R. Terry, 1994, Outside directors and the adoption of poison pills, *Journal of Financial Economics* 35, 371-390.

Byrd, J., and K. Hickman, 1992, Do outside directors monitor managers: Evidence from tender offer bids. *Journal of Financial Economics* 32, 195-222.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross, 1989, Communication in the battle of the sexes game: Some experimental results, *RAND Journal of Economics* 20, 568-587.

Eckel, C., and C. Holt, 1989, Strategic voting in agenda-controlled committee experiments, *The American Economic Review* 79, 763-773.

Farrell, J., and M. Rabin, 1996, Cheap talk, *Journal of Economic Perspectives* 10, 103-118.

Gilson, R., and R. Kraakman, 1991, Reinventing the outside director: An agenda for institutional investors,*Stanford Law Review* 43, 863-906.

Hermalin, B., and M. Weisbach, 1991, *Endogenously chosen boards of directors and their monitoring of the CEO*, *American Economic Review* 88, 96-118.

Isaac, M., and J. Walker, 1988, Communication and free-riding behavior: The voluntary contribution mechanism,*Economic Inquiry* 26, 585-608.

Jensen, M., 1993, The modern industrial revolution, exit, and the failure of internal control systems. *Journal of Finance* 48, 831-880.

Milgrom, P., and J. Roberts, 1996, *Coalition-proofness and correlation with arbitrary communication possibilities*, *Games and Economic Behavior* 17, 113-28.

Noe, T., and M. Rebello, 1997, The design of corporate boards: Composition, compensation, factions, and turnover, working Paper, Georgia State University.

Palfrey, T., 1990, Implementation in Bayesian equilibrium: The multiple equilibrium problem in mechanism design, in *Advances in Economic Theory*. J-J Laffont ed., Cambridge: Cambridge University Press.

Palfrey, T., and S. Srivastava, 1993, *Bayesian implementation*, Chur, Switzerland: Harwood Academic Publishers.

Rosenstein, S., and J. Wyatt, 1997, Inside directors, board effectiveness, and shareholder wealth, *Journal of Financial Economics* 44, 229-250.

Sefton, M., and A.Yavas, 1996, Abreu-Matsushima mechanisms: Experimental evidence, *Games and Economic Behavior* 16, 280-302.

Shivdasani, A., 1993, Board composition, ownership structure, and hostile takeovers, *Journal of Accounting and Economics* 16, 167-198.

Shleifer, A., and R. Vishny, 1989, Management entrenchment: The case of manager-specific investments, *Journal of Financial Economics* 25, 123-139.

Warther, V., 1998, Board effectiveness and board dissent: A model of the board's relationship to management and shareholders, *Journal of Corporate Finance* 4, 53-70.

Weisbach, M., 1988, Outside directors and CEO turnover, *Journal of Financial Economics* 20, 431-460.

Yermack, D., 1996, Higher market valuation of companies with a small board of directors, *Journal of Financial Economics* 40, 185-212.

| Treatment | **RANW** | **RAP** | **RANP** | **REP** | **RENP** |
|---|---|---|---|---|---|
| Mixing protocol | Random | Random | Random | Repeated | Repeated |
| Penalty for split votes | Yes | Yes | No | Yes | No |
| Number of watchdogs | 0 | 4 | 4 | 4 | 4 |
| Members in each group | 7 | 7 | 7 | 7 | 7 |
| Communication permitted | Yes | Yes | Yes | Yes | Yes |
| Mode of communication | Verbal | Verbal | Verbal | Verbal | Verbal |
| Communication protocol before project-quality draw | Group-subgroup | Group-subgroup | Group-subgroup | Group-subgroup | Group-subgroup |
| Number of groups | 3 | 3 | 2 | 3 | 2 |
| Number of draws | 30 | 30 | 20 | 30 | 20 |
| Number of good draws | 18 | 18 | 12 | 10 | 8 |
| Average payoff to insiders ($) | 10.31 | 9.30 | 9.30 | 7.62 | 8.08 |
| Average payoff to outsiders ($) | NA | 6.20 | 6.20 | 5.73 | 5.45 |

**Table 1.** *Description of the central treatments*. This table provides a description of the main features of the five treatments that are central to our experiment. It also provides information on the number of sessions of each treatment as well as the distribution of draws for each treatment. As the table indicates the treatments can be divided into two groups based on the mixing protocols employed—random mixing where group membership was changed after every round but subjects retained their agent-type, and repeated groups where group composition remained unchanged for the duration of the session. Two other features were varied. No watchdogs were included in the treatment RANW while all the other treatments called for 4 watchdogs to be included in each seven-member session. Insiders did not face the possibility of a penalty following split votes in treatments RANP and RENP. In all these sessions, all group members were allowed discussion time followed by discussion time within sub-group of insiders and watchdogs before the draw for project quality.

| Treatment | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Mixing protocol | Repeated | Repeated | Repeated | Repeated | Repeated | Repeated | Random watchdogs | Random watchdogs | Random watchdogs |
| Penalty for split votes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Number of watchdogs | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| Members in each group | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |
| # of periods with communication | 0 | First 5 | First 5 | First 5 | First 5 | First 5 | First 5 | First 5 | First 5 |
| Mode of communication | NA | Verbal | Message | Message | Message | Message | Message | Message | Message |
| Communication protocol before project-quality draw | NA | Group-subgroup | Group-subgroup | Subgroup-group | Group only | Subgroup-group-subgroup | Group-subgroup | Subgroup-group | Group only |
| Number of groups | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| Number of draws | 30 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 |
| Number of good draws | 16 | 5 | 6 | 6 | 6 | 6 | 6 | 6 | 6 |
| Average payoff to insiders ($) | 7.63 | 8.75 | 8.50 | 8.75 | 9.30 | 8.20 | 8.75 | 8.75 | 8.75 |
| Average payoff to outsiders ($) | 4.93 | 6.00 | 6.00 | 6.00 | 6.20 | 5.80 | 6.00 | 6.00 | 6.00 |

**Table 2**. *Description of robustness sessions.* This table describes the key features of 11 robustness sessions. It also provides information on the number of sessions of each treatment as well as the distribution of draws for each treatment. As the table indicates the treatments can be divided into two groups based on the mixing protocols employed--repeated groups where group composition remained unchanged for the duration of the session, and random watchdogs where watchdogs were shuffled across groups after every round but insiders remained with their groups. Two other features were varied. No communication was permitted in three sessions while the remaining sessions only allowed for communication during the first five rounds. With one exception, communication after the draw for project quality took the form of insiders passing watchdogs a piece of paper after having circled either "yes" or "no" on it. In the one exception verbal communication was permitted between watchdogs and insiders. All communication between subjects prior to the project quality draw was verbal. The sequence of discussion periods for groups and sub-groups varied as presented in row 7.

| | All Draws | Good Draws | Bad Draws |
|---|---|---|---|
| *Random mixing* | | | |
| No watchdogs | 17<br>(30) | 17<br>(18) | 0<br>(12) |
| Watchdogs and penalty | 30<br>(30) | 18<br>(18) | 12<br>(12) |
| Watchdogs no penalty | 20<br>(20) | 12<br>(12) | 8<br>(8) |
| *Repeated groups* | | | |
| Watchdogs and penalty | 25<br>(30) | 6<br>(10) | 19<br>(20) |
| Watchdogs no penalty | 18<br>(20) | 7<br>(8) | 11<br>(12) |

**Table 3.** *Adoption of the efficient policy in the central treatments.* This table presents the frequency with which majority votes resulted in the adoption of the efficient policy for all the central treatments by type of draw. The number of draws are presented in parentheses. The efficient policy calls for the project to be accepted when it is good and rejected if it is bad.

| | Random no watchdogs | Random penalty | Random no penalty | Repeated penalty | Repeated no penalty |
|---|---|---|---|---|---|
| Random no watchdogs | | | | | |
| Random penalty | 1.02<br>(24) | | | | |
| Random no penalty | 0.697<br>(15.68)* | UD†<br>(UD) | | | |
| Repeated penalty | 5.168**<br>(28.48)* | 8.33*<br>(0.62) | 5.857**<br>(0.412) | | |
| Repeated no penalty | 0.376<br>(20.32)* | 2.367<br>(1.035) | 2.78***<br>(0.70) | 1.67<br>(0.14) | |

† Undefined as the two distributions are identical.
Note: $\chi^2_{(1,.10)} = 2.71$, $\chi^2_{(1,.05)} = 3.84$, and $\chi^2_{(1,.01)} = 6.63$.

**Table 4.** *Tests for differences in incedence of the adoption of the efficient policy in the central treatments.* This table presents Chi-squares values for the distribution of efficient votes across experimental treatments. For each test, the adopted policy is categorized as either efficient or inefficient. An adopted policy is classified as being an efficient if the majority vote results in the project being accepted when it is good and rejected if it is bad. Otherwise it is classified as inefficient. Each cell contains two values. The first statistic in each cell represents the Chi-square statistic for distributions of votes following good draws. The second number which appears in parentheses represents the Chi-square statistic for vote distributions following bad draws. Significance at the 10%, 5%, and 1% confidence levels is denoted by ***, **, and * respectively.

| | All yes | 1 No | 2 No | All No | Total draws |
|---|---|---|---|---|---|
| *Random mixing* | | | | | |
| No watchdogs | 16<br>(12) | 1<br>(0) | 0<br>(0) | 1<br>(0) | 18<br>(12) |
| Watchdogs and penalty | 18<br>(0) | 0<br>(0) | 0<br>(0) | 0<br>(12) | 18<br>(12) |
| Watchdogs and no penalty | 12<br>(0) | 0<br>(0) | 0<br>(0) | 0<br>(0) | 12<br>(8) |
| *Repeated groups* | | | | | |
| Watchdogs and penalty | 7<br>(3) | 0<br>(1) | 1<br>(0) | 2<br>(16) | 10<br>(20) |
| Watchdogs and no penalty | 8<br>(5) | 0<br>(0) | 0<br>(0) | 0<br>(7) | 8<br>(12) |

**Table 5.** *Aggregate insider voting patterns in the central treatments.* This figure presents insider vote distributions in the central treatments. The vote distributions are classified by treatments. Each cell contains two numbers the first represents the frequency of votes following good draws. The second figure (in parentheses) is the frequency of votes following bad draws. The last column presents the total number of good draws and bad draws in each treatment.

| | Random penalty | Random no penalty | Repeated penalty | Repeated no penalty |
|---|---|---|---|---|
| Random penalty | | | | |
| Random no penalty | UD‡ (UD) ‡ | | | |
| Repeated penalty | 3.17† (2.74) † | 2.626† (1.868) † | | |
| Repeated no penalty | UD‡ (6.34) ††** | UD‡ (4.44) ††** | 2.88† (3.22) † | |

‡ The statistic is undefined as the two distributions are identical.
† A category of voting patterns was combined due to zero observations, degrees of freedom is 2.
†† Two categories of voting patterns was combined due to zero observations, degrees of freedom is 1.
$\varkappa^2_{(1,.10)} = 2.71$, $\varkappa^2_{(1,.05)} = 3.84$, and $\varkappa^2_{(1,.01)} = 6.63$.
$\varkappa^2_{(2,.10)} = 4.61$, $\varkappa^2_{(2,.05)} = 5.99$, and $\varkappa^2_{(2,.01)} = 9.21$.

**Table 6.** *Chi-square statistics for insider vote distributions in the central treatments including watchdogs.* Each cell contains two numbers. The first number represents the Chi-square statistic for vote distributions following good draws and the second number, in parentheses, represents the Chi-square statistic for insider vote distributions following bad draws. Significance at the 10%, 5%, and 1% confidence levels is denoted by ***, **, and * respectively.

| | 4-A; 0-N | 3-A; 1-N | 2-A; 2-N | 1-A; 3-N | 0-A; 4-N | Total draws |
|---|---|---|---|---|---|---|
| *Random mixing* | | | | | | |
| Watchdogs and penalty | 18 | 0 | 0 | 0 | 0 | 18 |
| | (0) | (0) | (1) | (0) | (11) | (12) |
| Watchdogs and no penalty | 12 | 0 | 0 | 0 | 0 | 12 |
| | (0) | (1) | (1) | (0) | (6) | (8) |
| *Repeated groups* | | | | | | |
| Watchdogs and penalty | 4 | 3 | 0 | 1 | 2 | 10 |
| | (5) | (1) | (2) | (0) | (12) | (20) |
| Watchdogs and no penalty | 7 | 0 | 0 | 0 | 1 | 8 |
| | (1) | (0) | (0) | (1) | (10) | (12) |

**Table 7.** *Watchdog vote distributions in the central treatmenst including watchdogs.* This table presents watchdog voting patterns by type of draw. Each cell presents the frequencies of a combination of abstain (A) and no (N) votes conditional on project quality. The first number in each cell presents the frequency of votes conditional on good draws and the second (in parentheses) presents the frequancy of votes conditional on bad draws. The total number of good and bad draws in each treatment appear in the last column.

| | Random penalty | Random no penalty | Repeated penalty | Repeated no penalty |
|---|---|---|---|---|
| Random penalty | | | | |
| Random no penalty | UD‡ (1.74) ††† | | | |
| Repeated penalty | 13.76††* (3.98) †† | 9.92††** (2.68) †† | | |
| Repeated no penalty | 2.36†††† (3.05) †† | 1.58†††† (4.38) † | 4.99†† (5.17) † | |

‡ Undefined as the two distributions are identical.
† The degrees of freedom is 4.
†† A category of voting patterns was combined due to zero observations, degrees of freedom is 3.
††† Two categories of voting patterns were combined due to zero observations, degrees of freedom is 2.
†††† Three categories of voting patterns were combined due to zero observations, degrees of freedom is 1.

Note that, $\chi^2_{(1,.10)} = 2.71$, $\chi^2_{(1,.05)} = 3.84$, and $\chi^2_{(1,.01)} = 6.63$,
$\chi^2_{(2,.10)} = 4.61$, $\chi^2_{(2,.05)} = 5.99$, and $\chi^2_{(2,.01)} = 9.21$,
$\chi^2_{(3,.10)} = 6.25$, $\chi^2_{(3,.05)} = 7.81$, and $\chi^2_{(3,.01)} = 11.34$,
$\chi^2_{(4,.10)} = 7.78$, $\chi^2_{(4,.05)} = 9.49$, and $\chi^2_{(4,.01)} = 13.28$.

**Table 8.** *Chi-square statistics for watchdog vote distributions in central treatments using watchdogs.* Each cell contains two numbers. The first number represents the Chi-square statistic for vote distributions following good draws and the second number, in parentheses, represents the Chi-square statistic for insider vote distributions following bad draws. Significance at the 10%, 5%, and 1% confidence levels is denoted by ***, **, and * respectively.

| | All Draws | Good Draws | Bad Draws | Total number of draws |
|---|---|---|---|---|
| *Random mixing* | | | | |
| No watchdogs | 16<br>(28) | 16<br>(16) | 0<br>(12) | 18<br>12 |
| Watchdogs and penalty | 30<br>(11) | 18<br>(0) | 12<br>(11) | 18<br>12 |
| Watchdogs no penalty | 20<br>(6) | 12<br>(0) | 8<br>(6) | 12<br>8 |
| *Repeated groups* | | | | |
| Watchdogs and penalty | 22<br>(11) | 6<br>(2) | 16<br>9 | 10<br>20 |
| Watchdogs no penalty | 14<br>(7) | 7<br>(0) | 7<br>(7) | 8<br>12 |

**Table 9.** *Frequency with which votes corresponded to vote vectors supporting equilibrium outcomes.* This table presents the frequency with which aggregate votes conformed with those supporting two competing equilibria—the efficient equilibrium and the coalition-proof equilibrium. Each cell in the second column presents the frequency with vote vectors conformed with those supporting efficient equilibria and coalition-proof equilibria (in parentheses). The next two columns presents this data following good and bad draws respectively. The final column first presents the number of good draws and then the number of bad draws for each treatment. In treatments containing watchdogs, efficient equilibria call for unanimous insider support following a good draw and rejection following a bad one. At least two watchdogs abstain following a good draw. In coalition-proof equilibria, all insiders and at least three watchdogs vote to reject at all times. In the treatments with only insiders, the coalition-proof equilibrium requires that all insiders vote to accept the project regardless of the draw.

| | All Draws | Good Draws | Bad Draws |
|---|---|---|---|
| | **Central treatment** | | |
| REP | 25/30 = 83.3% | 6/10 = 60% | 19/20 = 95% |
| | **Robustness treatments** | | |
| *No communication* | | | |
| RENC | 19/30 = 63.3% | 9/16 = 56.3% | 10/14 = 71.4% |
| *Face to face* | | | |
| Subgroup/group | 10/10 = 100% | 5/5 = 100% | 5/5 = 100% |
| *Message-no message* | | | |
| Group/subgroup | 9/10 = 90% | 5/6 = 83.3% | 4/4 = 100% |
| Subgroup/group | 9/10 = 90% | 5/6 = 83.3% | 4/4 = 100% |
| Group only | 10/10 = 100% | 6/6 = 100% | 4/4 = 100% |
| Subgroup/group/subgroup | 8/10 = 80% | 4/6 = 66.7% | 4/4 = 100% |
| *Message-no message and random watchdogs* | | | |
| Group/subgroup | 9/10 = 90% | 5/6 = 83.3% | 4/4 = 100% |
| Subgroup/group | 9/10 = 90% | 5/6 = 83.3% | 4/4 = 100% |
| Group only | 9/10 = 90% | 5/6 = 83.3% | 4/4 = 100% |

**Table 10.** *Sensitivity of efficient majority votes to changes in mixing and communication protocols.* This table presents the percentage of efficient majority votes for all draws and by type of draw, for one central treatment and the robustness treatments. A vote is classified as being an efficient majority vote if the majority vote results in the project being accepted when it is good and rejected if it is bad. The central treatment presented is the *repeated groups with penalty* protocol (REP). The robustness treatments are classified first by the form of communication, then by the protocol governing the sequence of communication. With the exception of the last three treatments in the table, they all employed the *repeated groups* mixing protocol. In the final three treatments, the group of insiders remained together for all 10 draws while the watchdogs were randomly assigned to groups following each vote.

| | Good draws | Bad draws |
|---|---|---|
| *No communication* | | |
| RENC | 0.04 | (3.65)*** |
| *Face to face* | | |
| Subgroup/group | 2.73*** | (0.66) |
| *Message-no message* | | |
| Group/subgroup | 0.96 | (0.22) |
| Subgroup/group | 0.96 | (0.22) |
| Group only | 3.20*** | (0.22) |
| Subgroup/group/subgroup | 0.073 | (0.22) |
| *Message-no message and random watchdogs* | | |
| Group/subgroup | 0.96 | (0.22) |
| Subgroup/group | 0.96 | (0.22) |
| Group only | 0.96 | (0.22) |

**Table 11.** *Robustness of efficient outcome distributions* ***to changes in mixing and communication protocols***. In this table we present evidence on the robustness of the pattern of insider voting to changes in the communication and mixing protocols. The table contains Chi-square statistics comparing the distribution of efficient votes with the central treatment employing the *repeated groups with penalties* (REP) protocol. Statistics for outcome distributions conditioned on good draws are presented first while the statistics conditioned on bad draws are presented in parentheses. The robustness treatments are classified first by the form of communication, then by the protocol governing the sequence of communication. With the exception of the last treatment in the table, they all employed the *repeated groups* mixing protocol. In this treatment, the group of insiders remained together for all 10 draws while the watchdogs were randomly assigned to groups following each vote. Significance at the 10%, 5%, and 1% confidence levels is denoted by ***, **, and * respectively.

| | All Draws | Good Draws | Bad Draws | Total number |
|---|---|---|---|---|
| *No communication* | | | | |
| RENC | 10 | 9 | 1 | 16 |
| | (0) | (0) | (0) | 14 |
| *Face to face* | | | | |
| Subgroup/group | 10 | 5 | 5 | 5 |
| | (5) | (0) | (5) | 5 |
| *Message-no message* | | | | |
| Group/subgroup | 9 | 5 | 4 | 6 |
| | (4) | (0) | (4) | 4 |
| Subgroup/group | 9 | 5 | 4 | 6 |
| | (4) | (1) | (3) | 4 |
| Group only | 10 | 6 | 4 | 6 |
| | (4) | (0) | (4) | 4 |
| Subgroup/group/subgroup | 8 | 4 | 4 | 6 |
| | (3) | (2) | (1) | 4 |
| *Message-no message and random watchdogs* | | | | |
| Group/subgroup | 8 | 5 | 3 | 6 |
| | (2) | (1) | (1) | 4 |
| Subgroup/group | 9 | 5 | 4 | 6 |
| | (3) | (0) | (3) | 4 |
| Group only | 9 | 5 | 4 | 6 |
| | (3) | (1) | (2) | 4 |

**Table 12.** *Frequency with which votes in the robustness treatments corresponded to vote vectors supporting equilibrium outcomes.* This table presents the frequency with which aggregate votes conformed with those supporting two competing equilibria—the efficient equilibrium and the coalition-proof equilibrium. Each cell in the second column presents the frequency with vote vectors conformed with those supporting efficient equilibria and coalition-proof equilibria (in parentheses). The next two columns presents this data following good and bad draws respectively. The final column first presents the number of good draws and then the number of bad draws for each treatment. Efficient equilibria call for unanimous insider support following a good draw and rejection following a bad one. At least two watchdogs abstain following a good draw. In coalition-proof equilibria, all insiders and at least three watchdogs vote to reject at all times.

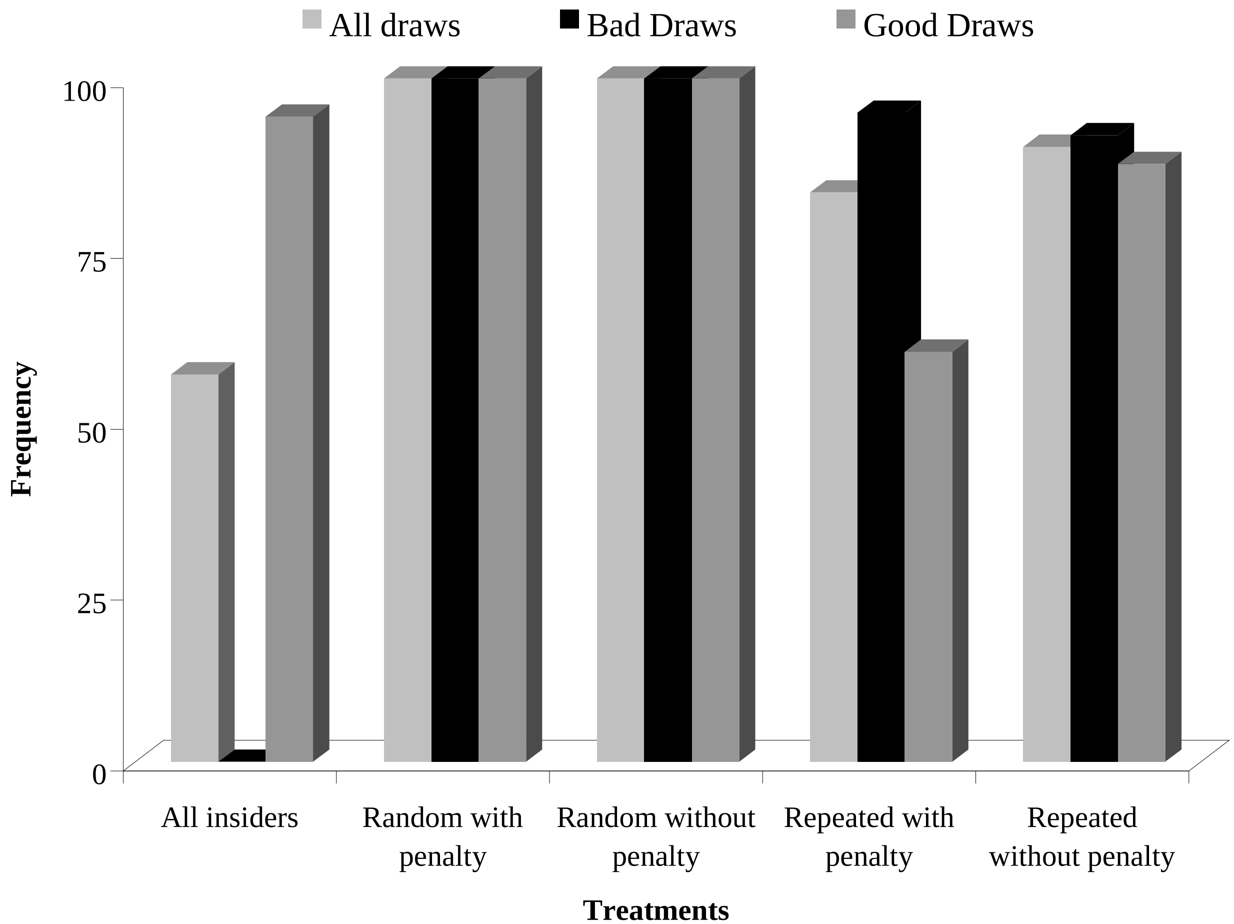


**Figure 1.** *Incidence of institutionally-preferred outcomes in the central treatments.* This figure presents the proportion of institutionally-preferred outcomes for all the central treatments. This information is presented for all draws in the treatment and then by type of draw. The lightest shaded columns represent the percentage of institutionally-preferred outcomes following for each treatment. The black columns represent the percentage of institutionally-preferred outcomes following bad draws and while the darker gray columns represent the percentage of institutionally-preferred outcomes following good draws. An outcome is classified as being institutionally-preferred if the majority vote results in the project being accepted when it is good and rejected if it is bad.

**Figure 2. Panel A.** *Treatment REP*.

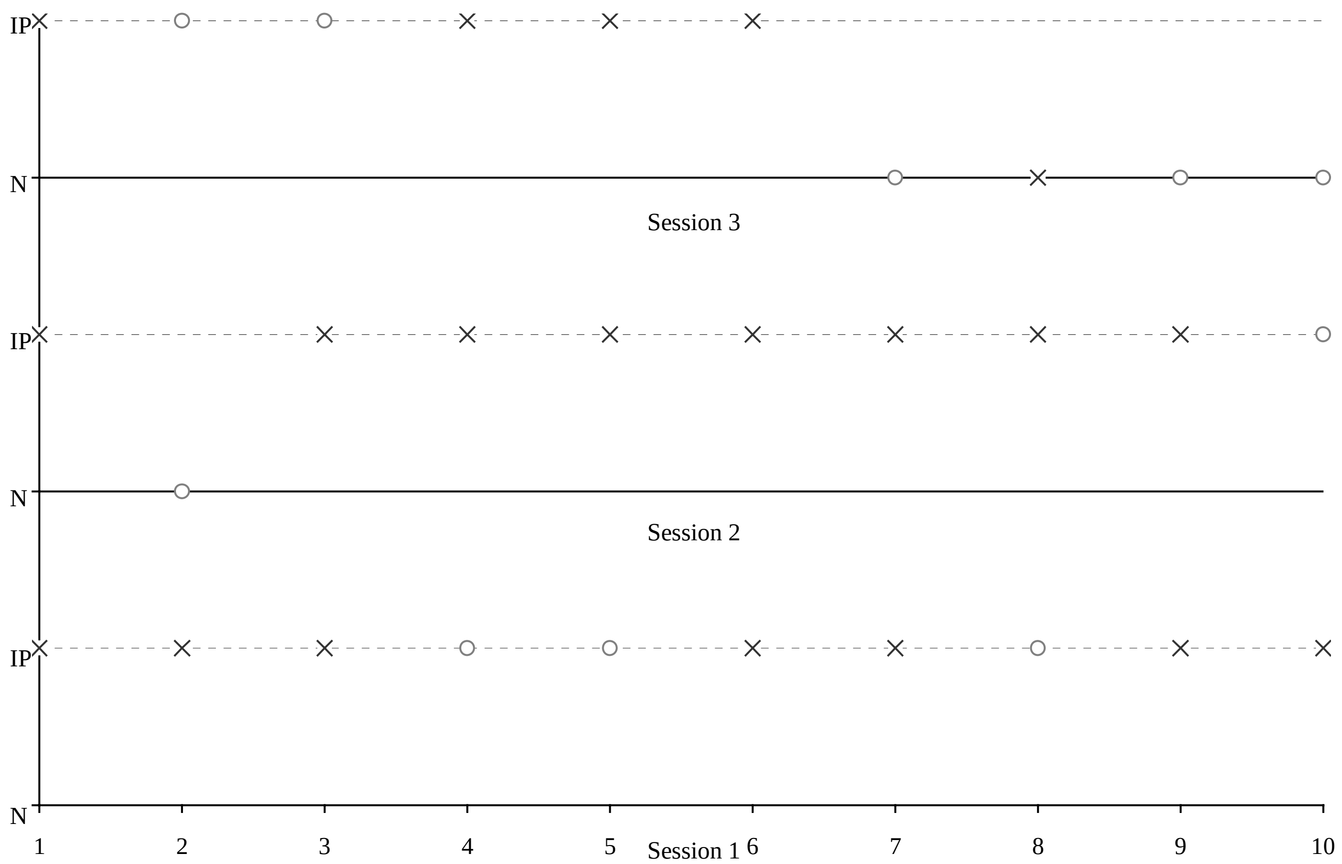

**Figure 2. Panel B.** *Treatment RENP.*

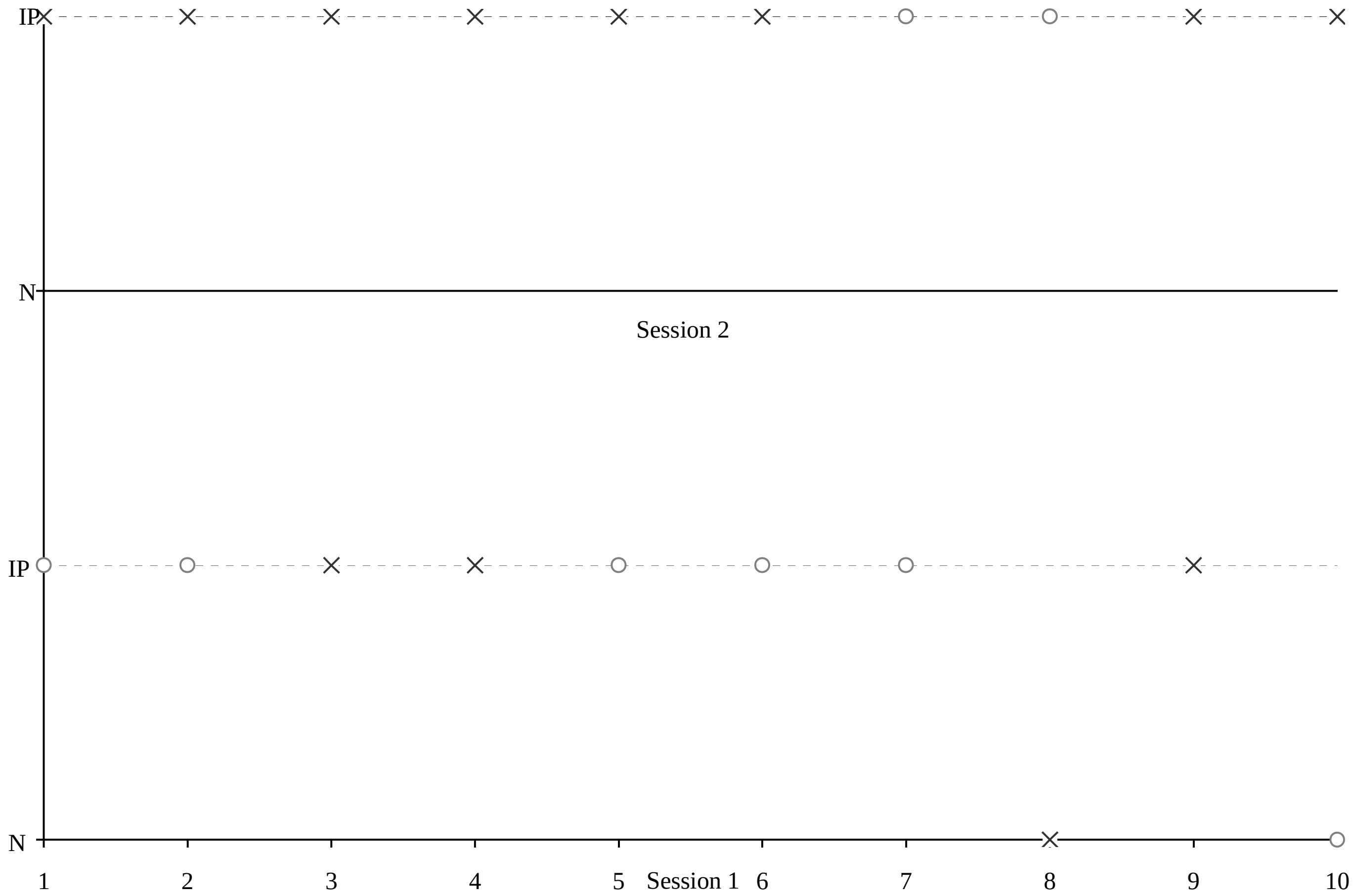


**Figure 2.** *Outcomes by period in the repeated treatments.* This figure presents the outcome of each vote in two central treatments: Panel A illustrates outcomes for the repeated groups with penalties (REP) treatment and Panel B for the repeated groups without penalties treatment (RENP). In each case, the X-axis represents the number of periods. Outcomes from the three sessions in this treatment are stacked one on top of the other. Each outcome is classified as institutionally-preferred (IP) or not (N). Outcomes following bad draws are represented by X, while outcomes following good draws are represented by O.

**Figure 3. Panel A.** *Good draws.*

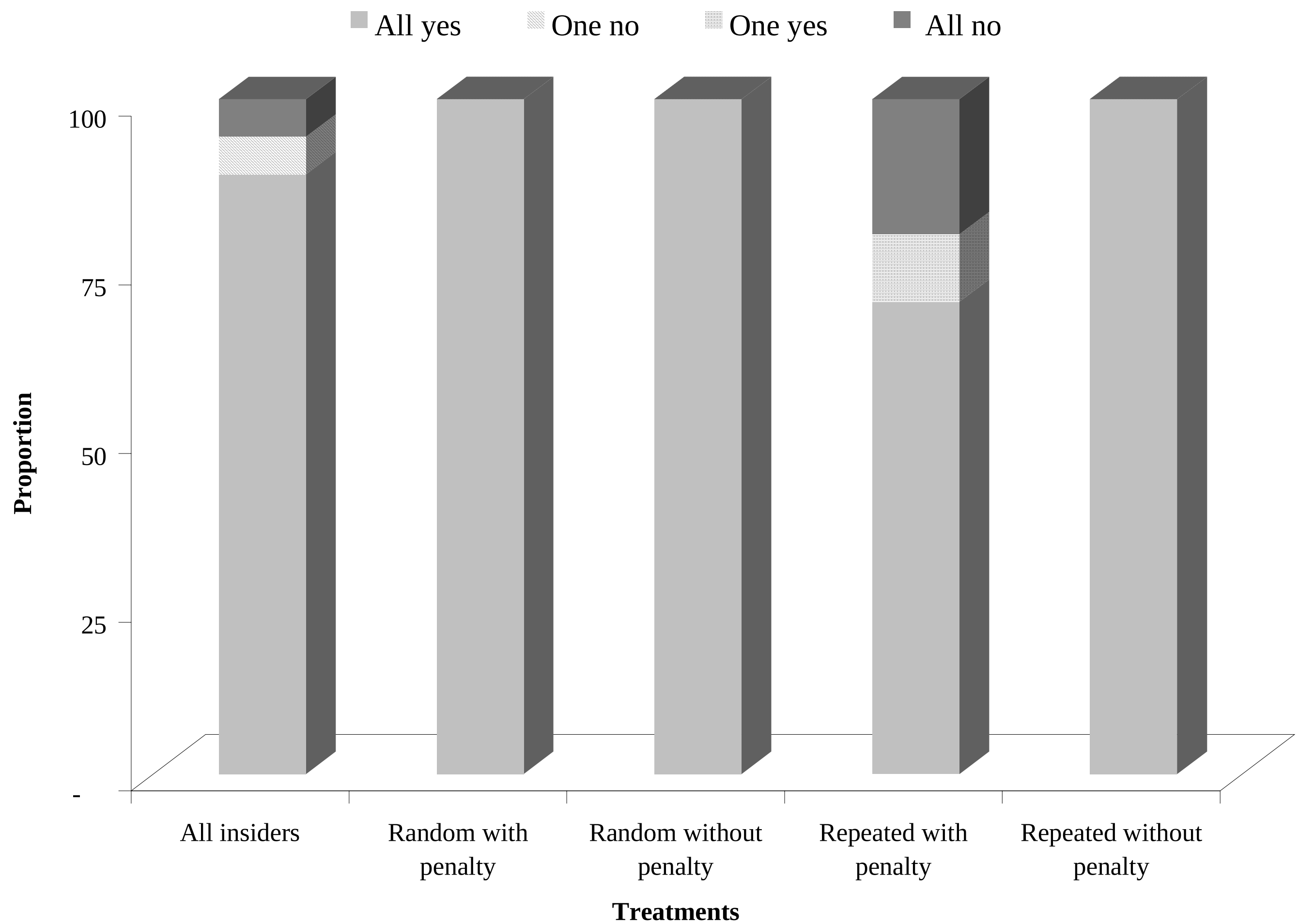

**Figure 3. Panel B.** *Bad draws*

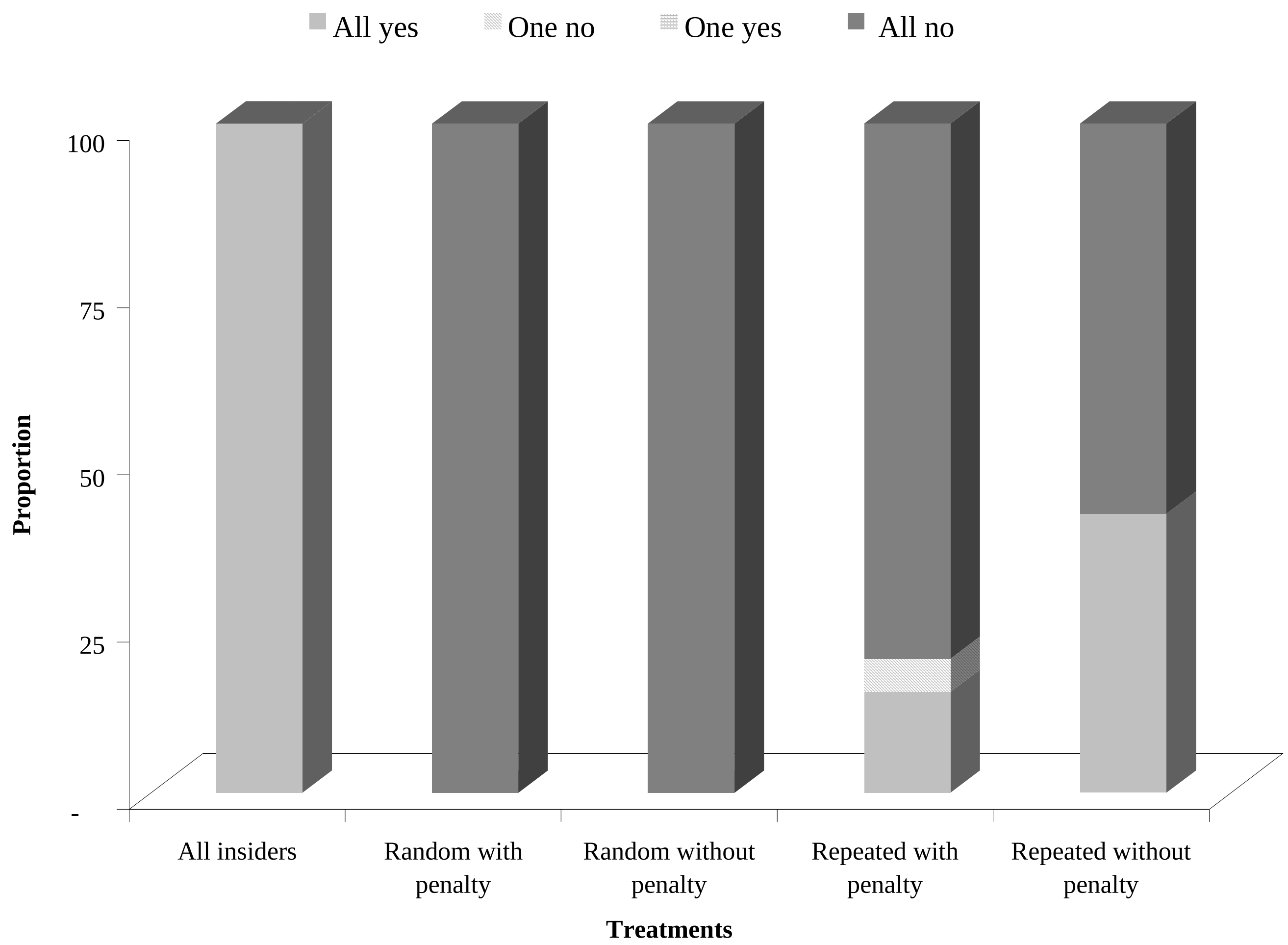


**Figure 3.** *Aggregate insider voting patterns in the central treatments.* This figure presents insider voting patterns. Panel A presents voting patters following *good draws* and Panel B presents voting patterns following *bad* draws. The lightest shaded areas indicate the frequency with which insiders unanimously voted to accept the project. The area with diagonal hatches indicate the frequency with which two of three insiders in a group voted to accept the project. The dark gray areas represent unanimous votes to reject the project. The bricked areas represents two no votes.

**Figure 4. Panel A.** *Treatment REP*.

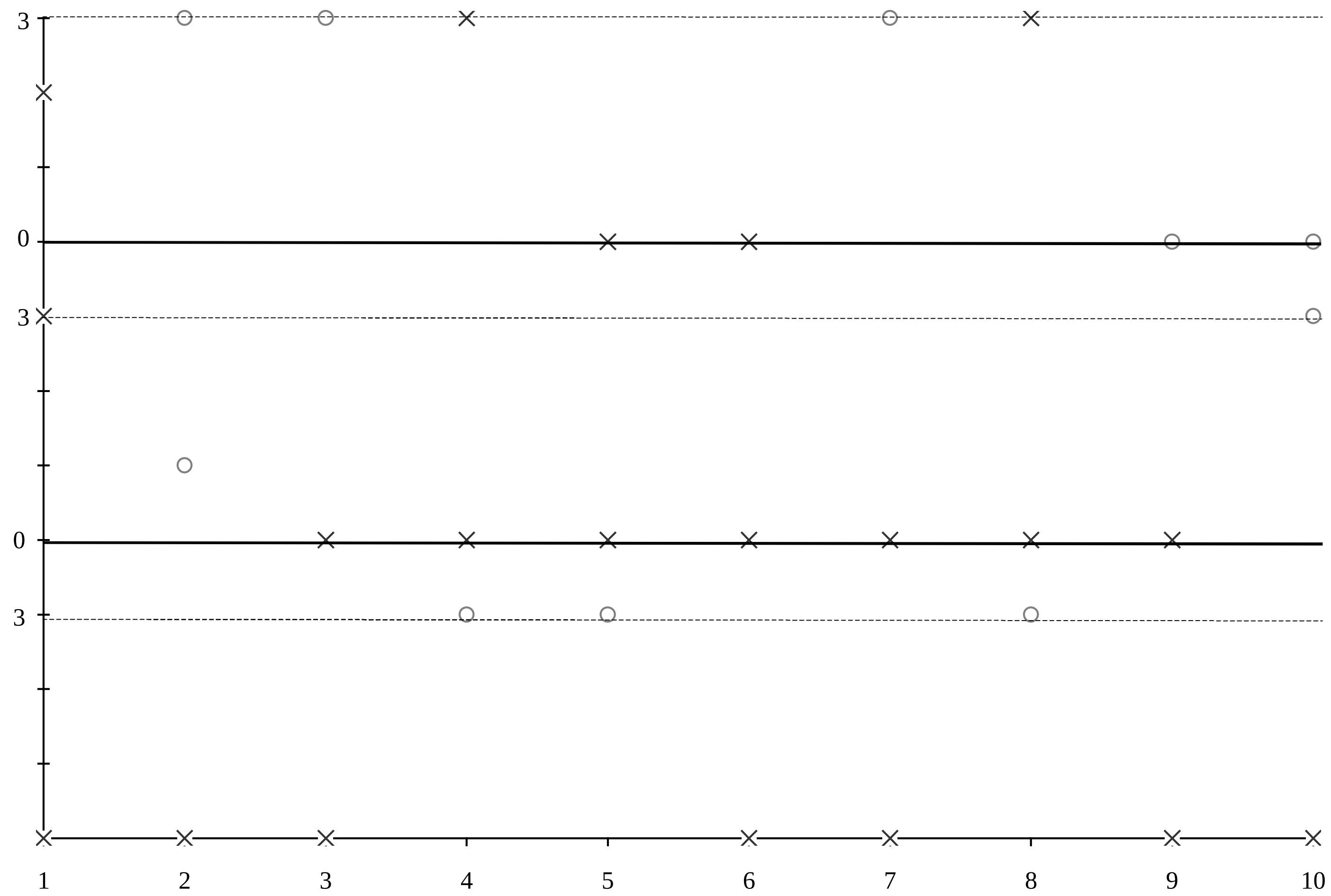

**Figure 4. Panel B.** *Treatment RENP.*

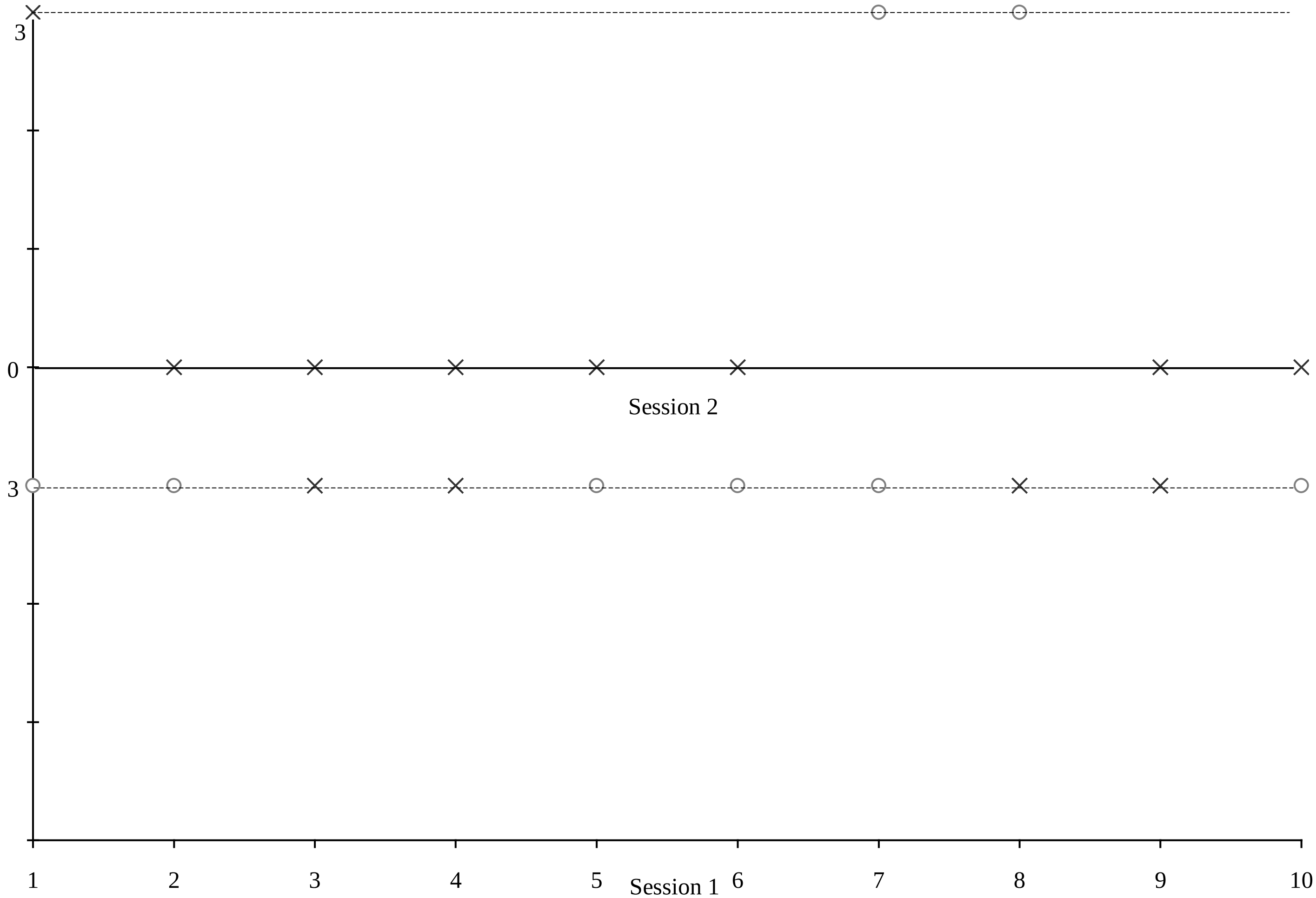


**Figure 4.** *Insider voting by periods.* This figure presents insider votes during each period of two core treatments. Panel A presents insider votes in the repeated group with penalties treatments (REP) while Panel B presents insider votes in the repeated groups without penalty treatment (RENP). In each case, the X-axis represents the number of periods. We graph the number of insider yes votes for each period. Votes following bad draws are represented by X, while votes following good draws are represented by O.

**Figure 5. Panel A.** *Good draws.*

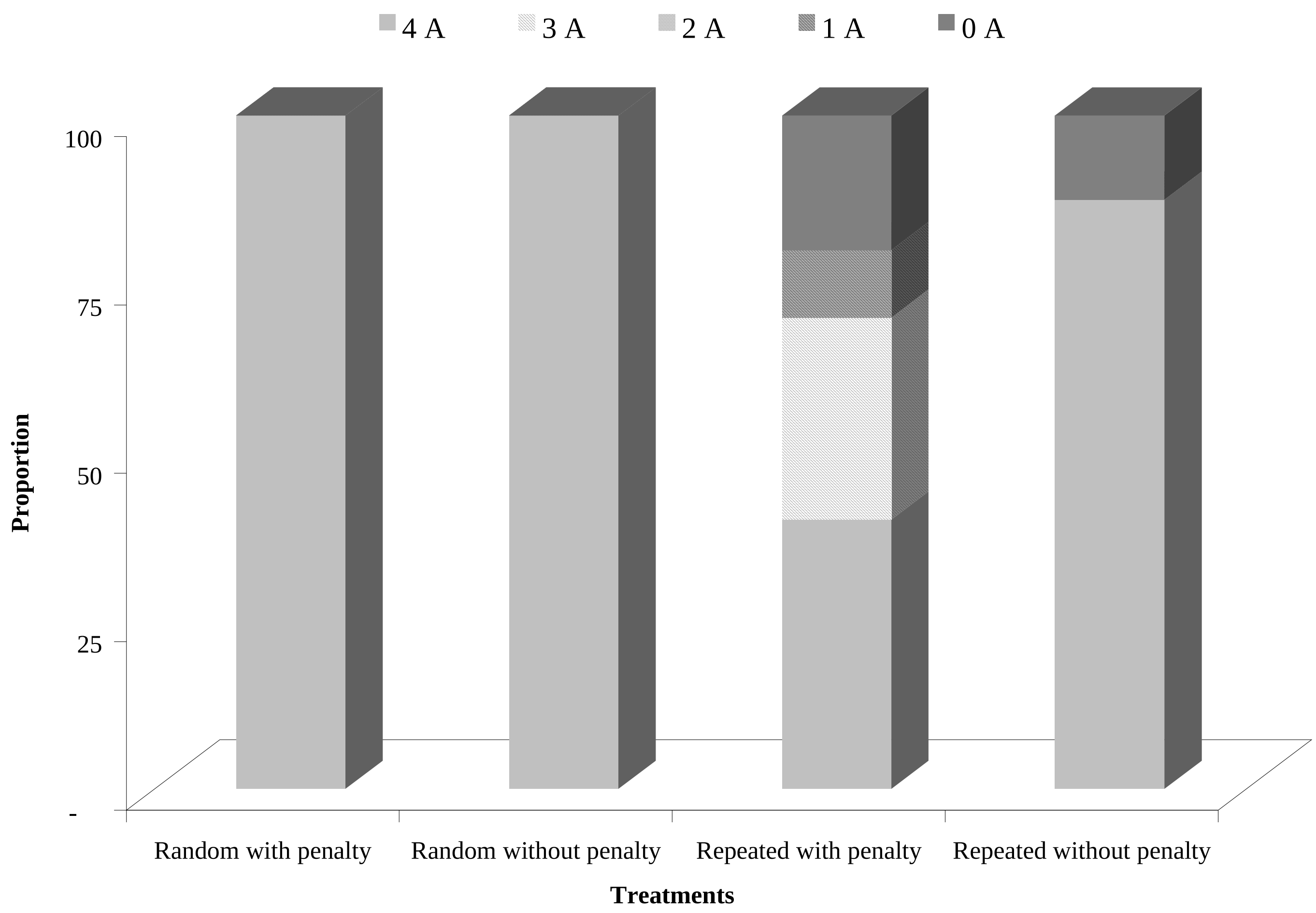

**Figure 5. Panel B.** *Bad draws*

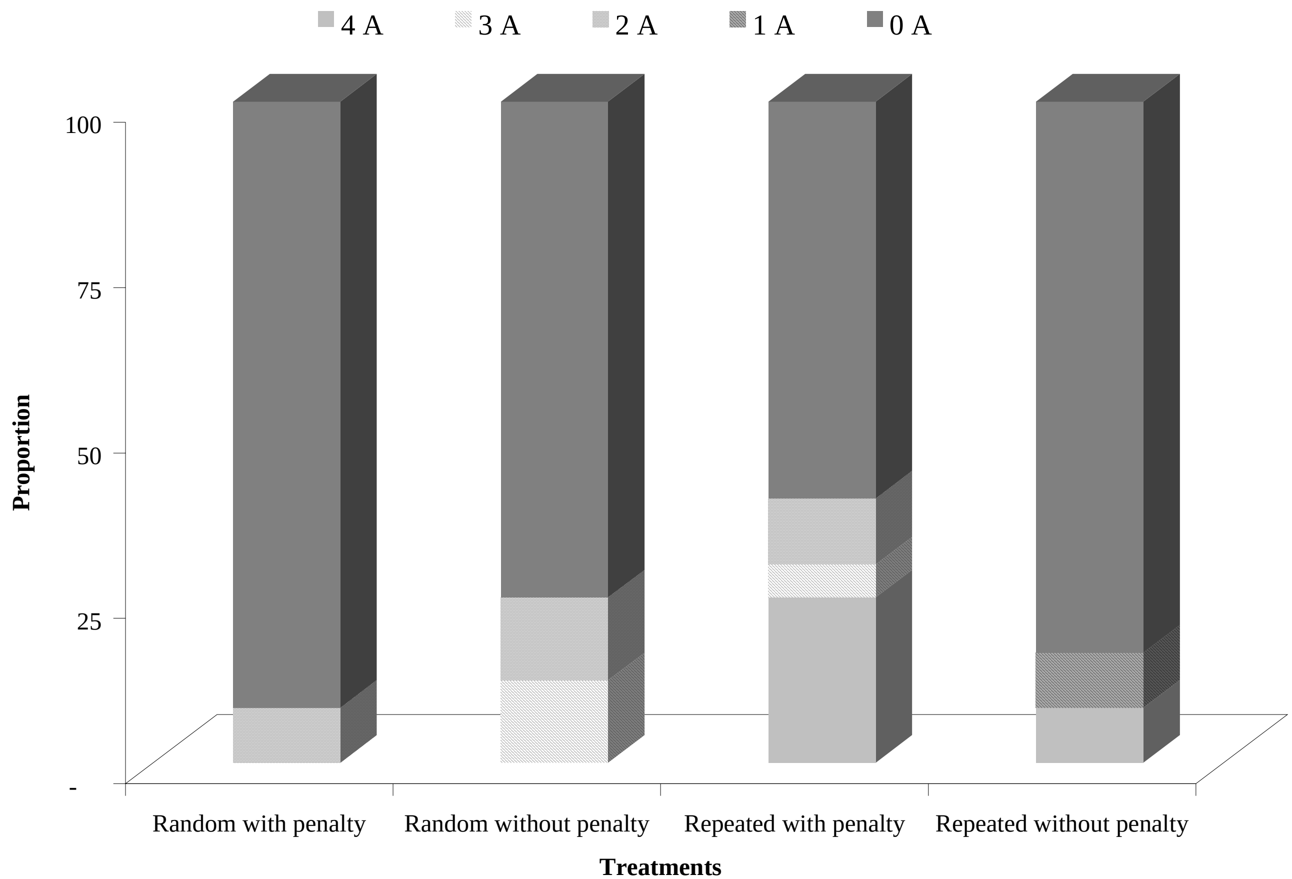


**Figure 5.** *Aggregate watchdog voting patterns.* This figure presents watchdog voting patterns. Panel A presents voting patterns following *good draws* and Panel B presents voting patterns following *bad* draws. The lightest shaded areas indicate the frequency with which watchdogs unanimously abstained. The light hatched areas indicate the frequency with which three of four watchdogs in a group abstained. The darkest areas represent unanimous votes to reject the project. The dark left-to-right hatched areas represents three no votes while the dark right-to-left hatched areas represent two no votes.

**Figure 6. Panel A.** *Treatment REP.*

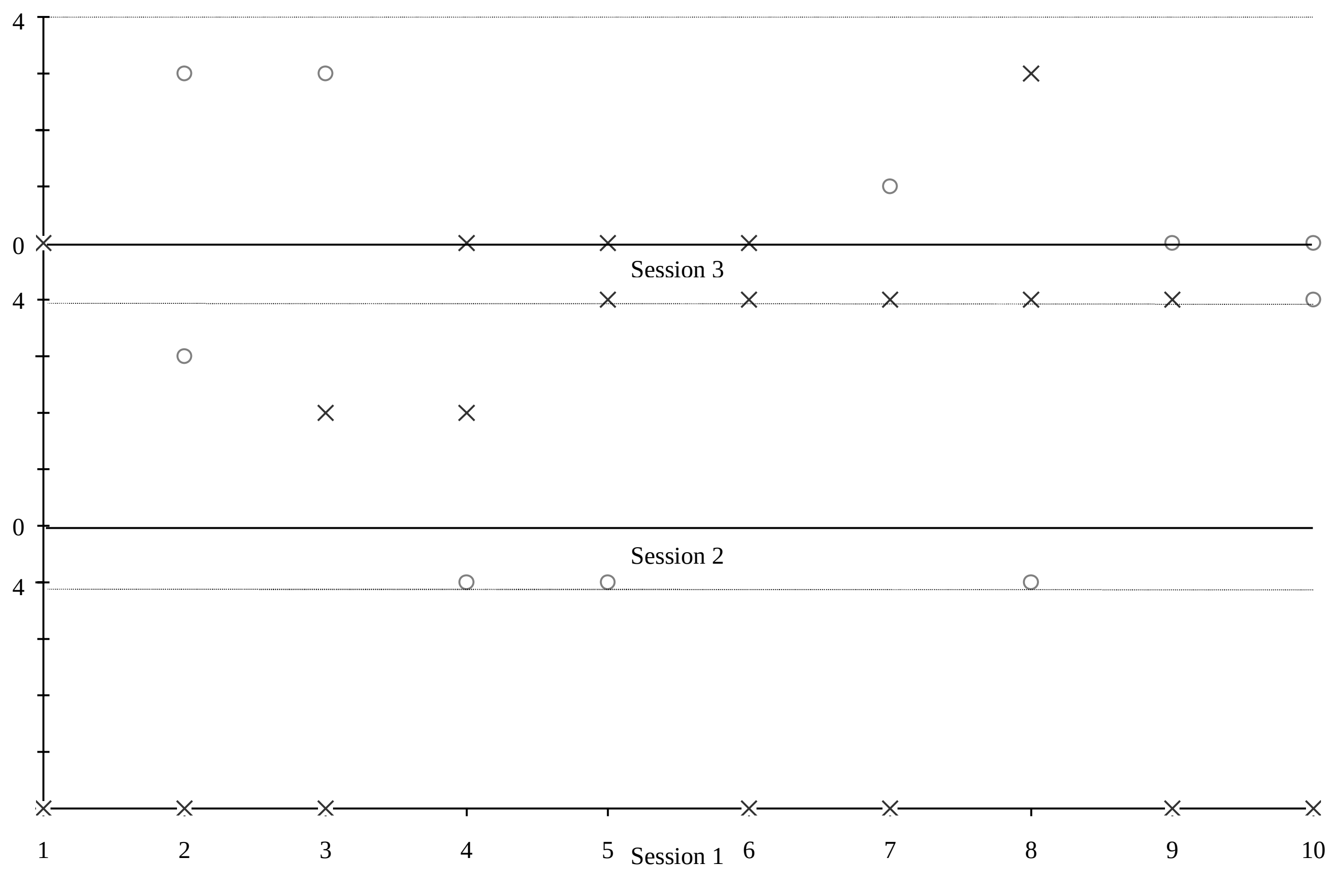

**Figure 6. Panel B.** *Treatment RENP.*

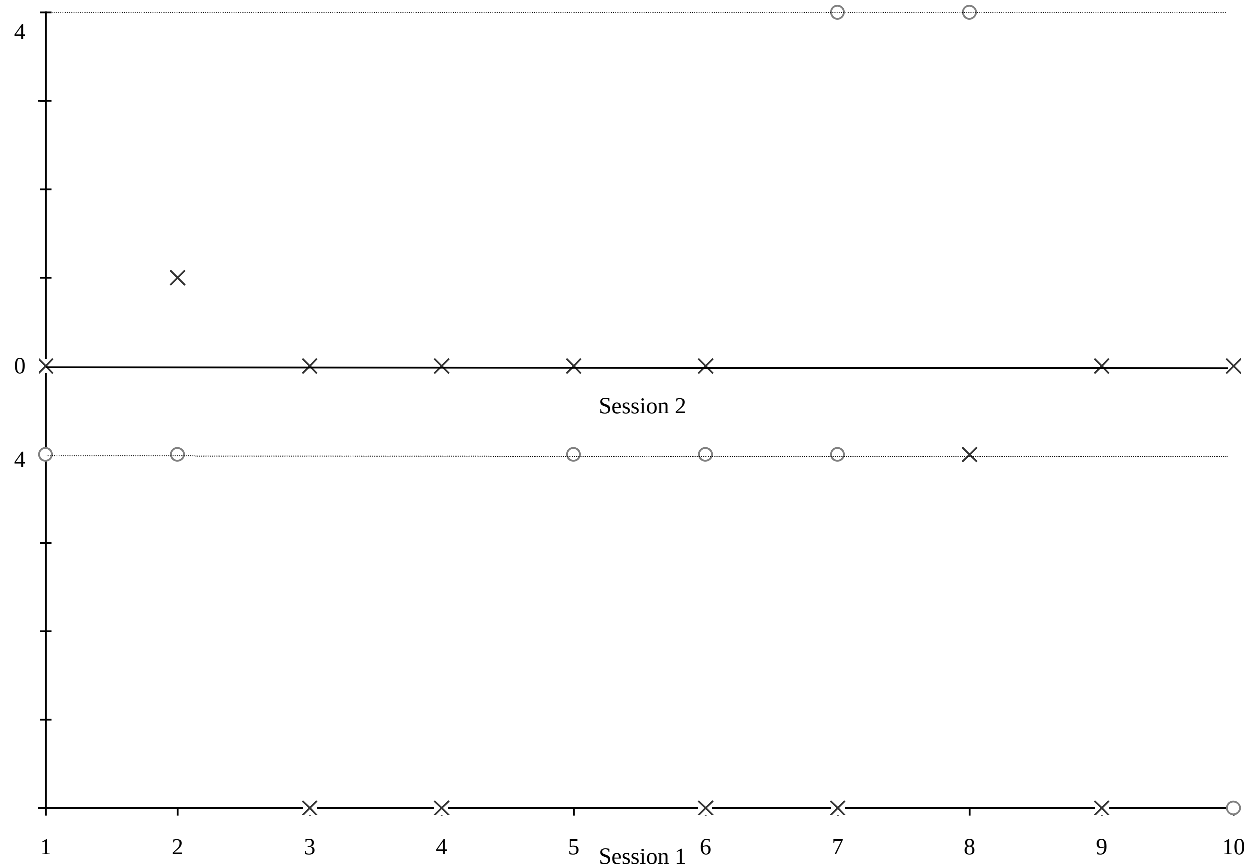


**Figure 6.** *Watchdog voting by periods.* This figure presents watchdog votes during each period of two core treatments. Panel A presents watchdog votes in the repeated group with penalties treatments (REP) while Panel B presents watchdog votes in the repeated groups without penalty treatment (RENP). In each case, the X-axis represents the number of periods. We graph the number of watchdog abstain votes for each period. Votes following bad draws are represented by X, while votes following good draws are represented by O.

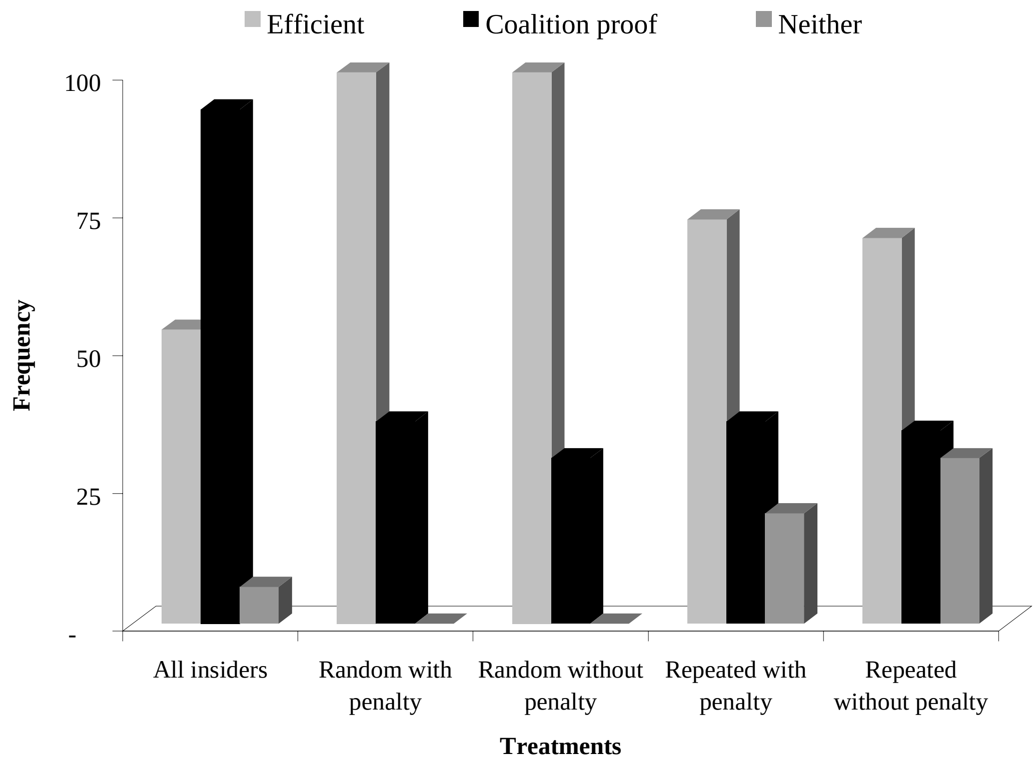


**Figure 7.** *Frequency with which votes corresponded to vote vectors supporting equilibrium outcomes.* This figure graphs the frequency with which aggregate votes conformed with those supporting two competing equilibria—the efficient equilibria and the coalition-proof equilibria. It also presents the percentage of times subject votes do not correspond to the vote vectors supporting these two equilibria. In treatments containing watchdogs, efficient equilibria call for unanimous insider support following a good draw and rejection following a bad one. At least two watchdogs abstain following a good draw. In coalition-proof equilibria, all insiders vote and at least three watchdogs vote to reject at all times. In the treatments with only insiders, the coalition-proof equilibrium requires that all insiders vote to accept the project regardless of the draw.

**Figure 8. Panel A.** *Good draws.*

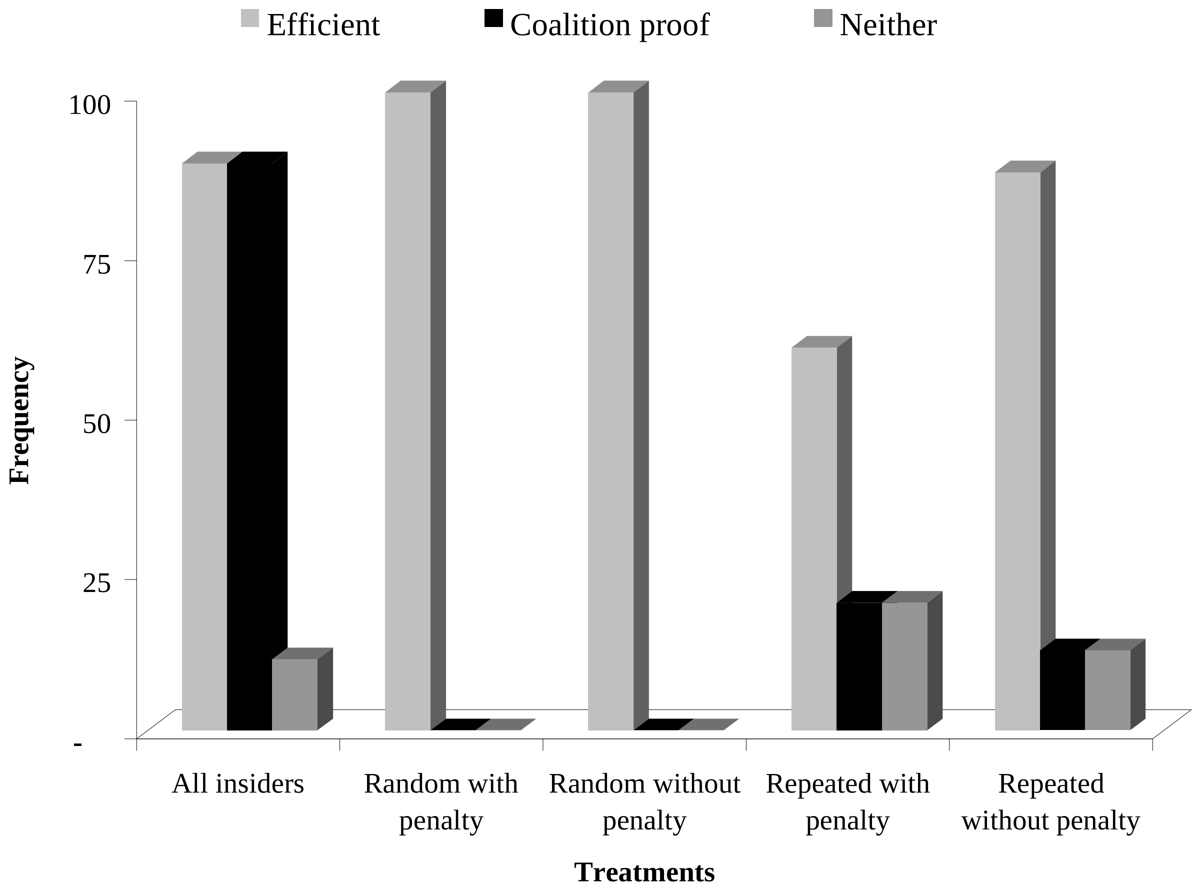

**Figure 8. Panel B.** *Bad draws*

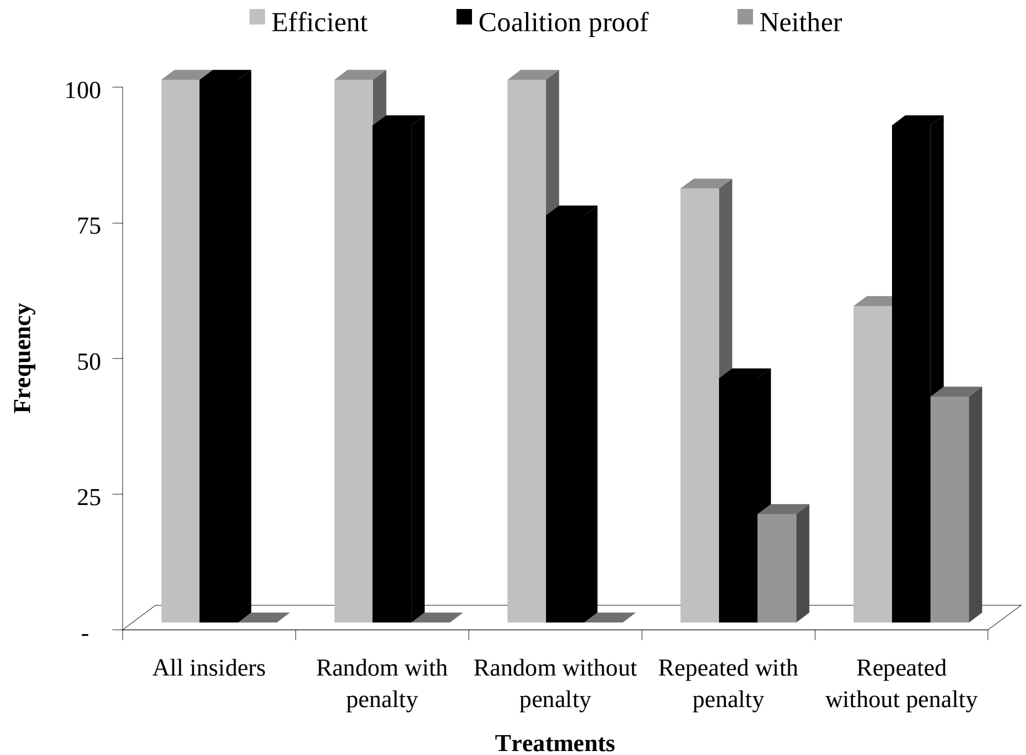


**Figure 8.** *Frequency with which votes corresponded to vote vectors supporting equilibrium outcomes conditioned on draws.* This figure graphs the frequency with which aggregate votes conformed with those supporting two competing equilibria—the efficient equilibria and the coalition-proof equilibria. It also presents the percentage of times subject votes do not correspond to the vote vectors supporting these two equilibria. Panel A (B) presents this information conditioned on good (bad) draws. In treatments containing watchdogs, efficient equilibria call for unanimous insider support following a good draw and rejection following a bad one. At least two watchdogs abstain following a good draw. In coalition-proof equilibria, all insiders vote and at least three watchdogs vote to reject at all times. In the treatments with only insiders, the coalition-proof equilibrium requires that all insiders vote to accept the project regardless of the draw.