

# The Pivotal Mechanism Revisited: Some Evidence on Group Manipulation

Francesco Feri\*, Anita Gantner<sup>†‡</sup>, Wolfgang Höchtl<sup>†</sup>  
and Rupert Sausgruber<sup>†</sup>

April 23, 2012

## Abstract

This paper studies the vulnerability of the pivotal mechanism with respect to manipulation by groups. In a lab experiment, groups decide on the implementation of various alternatives, some of which imply opposite interests for the two subgroups. We investigate the occurrence of tacit and explicit collusion by allowing for communication within subgroups in one treatment and prohibiting it in another. Even though all agents' preferences are common knowledge and there exists a simple symmetric collusive strategy for one subgroup, we find little evidence for tacit collusion. Only when explicit communication is allowed, collusion is established. A behavioral model using quantal response equilibrium where we assume that subjects have beliefs over the correlation of errors of same-type subjects helps explain the main features of our data.

Keywords: Collective Decision Making, Pivotal Mechanism, Collusion, Communication, QRE

JEL Classification: D71, D61, C92

---

\*University of Trieste

†University of Innsbruck

‡Corresponding author: anita.gantner@uibk.ac.at, Universitätsstr.15, A-6020 Innsbruck, Tel.: +43/512/507/7377, Fax: +43/512/507/2980. We gratefully acknowledge financial support by the Austrian Science Fund (FWF) through Project No. S10307-G14 and by the University of Innsbruck through grant “Nachwuchsförderung”. Thanks to Karl Seiringer for excellent research assistance.

# 1 Introduction

The pivotal mechanism (Clarke 1971) is a demand revealing mechanism developed for public goods decisions which proposes an efficient solution to the question of whether or not a public good of a given size should be provided. It uses a transfer system such that each individual takes into account the marginal social impact on the rest of society made by this individual's vote or report. In environments with quasi-linear preferences, revelation of true preferences is then a dominant strategy.<sup>1</sup> The pivotal mechanism has a remarkable standing in the literature: It is one of the most well-known mechanisms in social choice theory, present in each textbook of public choice since decades. Despite being a nice theoretical construct, it has experienced a less successful history with respect to its applications in the context of public goods provision.

This paper reports the results of an experiment that studies how the lack of coalition-proofness affects the performance of the pivotal mechanism. Already in the 70s, Groves and Ledyard (1977a, 1977b) pointed out some shortcomings of demand revealing mechanisms that may affect the desired efficient allocation in a public goods context. Besides the failure to produce a budget balance they also noted that the mechanisms are vulnerable to strategic manipulation by coalitions of agents which would move the outcome away from Pareto optimality. Tideman and Tullock (1977), on the other hand, considered this a characteristic of all group processes including markets and majority rules, thus defending the applicability of demand-revealing processes. Green and Laffont (1979) later provided a formal proof for the impossibility of finding coalition-proof incentive compatible Groves mechanisms, even if only a single coalition with two or more agents can be formed.

Whether and when groups can successfully manipulate the pivotal mechanism is an empirical question, which will be studied in this paper. The possibility of manipulation is relevant in a situation which allows a (sub)group of decision makers to coordinate their actions in order to obtain an outcome that is preferred by this group. We shall also refer to such actions as collusive behavior. One setting in which a collusive outcome may be expected is when communication amongst agents is possible. But from a theoretical perspective, collusion may also occur *tacitly* in the absence of explicit communication, in particular in a setting with rational and selfish agents who have complete information about other agents' preferences. Empirically, it is far from clear *a priori* whether and when collusion will actually occur in a given setup. There are at least three immediate motives for which

---

<sup>1</sup>Groves (1973), as well as Groves and Loeb (1975), and much earlier Vickrey (1961) independently also discovered such incentive compatible demand revealing mechanisms for environments with separable utility functions. Green and Laffont (1977) showed that the class of mechanisms proposed by Groves includes all these mechanisms, and furthermore that any efficient and strategy-proof direct revelation mechanism is isomorphic to a Groves mechanism.

such behavior may not be observed: First, a misrepresentation of preferences that leads to higher payoffs for a coalition requires that agents are able to identify such manipulation possibilities, thus limitations in cognitive abilities may prevent agents from collusive behavior. Second, agents may not be sufficiently confident about collusive behavior of other agents, thus complete information about other agents' preferences might not be sufficient to predict the occurrence of collusion as long as there is uncertainty about others' choices. And third, if one departs from the assumption of purely self-interested agents, it is possible that social preferences prevent agents from collusive behavior if the consequences are harsh for others. Expressed differently, the mechanism might work despite the theoretical prediction of collusion.

Experimental methods seem to be a useful tool to identify properties of this mechanism that are responsible for the difficulties in application. So far, experimental literature on the pivotal mechanism focussed on the (serious) problem that people do not seem to choose dominant strategies, that is, truth-telling is actually not implemented in the pivotal mechanism.<sup>2</sup> It was always suspected that this mechanism is too complex to be well-understood for applications and it therefore may not fulfill its intended purpose of revealing true preferences. This view was supported by the study of Attiyeh, Franciosi and Isaac (2000), who found that less than 10% of subjects revealed their true valuation for the public good, and moreover, no tendency toward the dominant strategy prediction was observed over time. The results suggested that due to its complexity the mechanism is inadequate for applications in the demand for public goods.<sup>3</sup> On the other hand, Kawagoe and Mori (2001) showed that there is a remedy to subjects' confusion caused by the complexity of the mechanism. When subjects were presented detailed payoff tables in addition to the abstract rule that maps bids to outcomes, nearly one half of the subjects played the dominant strategy. Kawagoe and Mori further argue that the bad performance of the pivotal mechanism in Attiyeh et al. (2000) may be due to the lack of *strict* incentive compatibility, i.e., since there exists a large number of strategies that leave subjects as well off as truth-telling for a wide range of strategies chosen by others, it is difficult for subjects to see why truth-telling is the unique dominant strategy. Finally, Cason et al. (2006) study the effect of secure implementation, which refers to mechanisms that ensure dominant strategy implementation with the additional requirement that there be no Nash equilibrium outcome other than the dominant strategy equilibrium outcome. This is not given in the pivotal mechanism, as there exist multiple Nash equilibria, in particular

---

<sup>2</sup>One of the early experiments by Tideman (1983) already suggested this possibility. However, this experiment was a kind of field experiment with little control over the factors that impact decisions, thus misrevelation could not be quantified.

<sup>3</sup>Note that in computer science applications, the class of Vickrey-Groves-Clarke mechanisms is widely used, e.g. in resolving task and resource allocation problems that occur in multi-agent systems (see e.g. Dash, Rogers and Jennings 2003, or Dash, Park and Jennings 2004).

bad Nash equilibria which are Pareto inferior to the dominant strategy outcome. Cason et al. (2006) found that the proportion of dominant strategy equilibrium outcomes increases from 50% for a pivotal mechanism, where implementation is not secure, to 81% for a securely implementable Groves-Clarke mechanism with single peaked preferences. As in Kawagoe and Mori (2001), the instructions contained detailed payoff tables. After the disappointing results of Attiyeh et al. (2000), these findings shed a much more positive light on the pivotal mechanism.

In this paper, we study the open question of the mechanism's vulnerability with respect to manipulation by groups. The existence of a simple misrepresentation of preferences that leads to an increase in payoffs does not necessarily imply that strategic manipulation is actually observed: This is the result of an experimental study on the Borda mechanism by Kube and Puppe (2009). They showed that the lack of strategy-proofness, which is a well-known flaw of this mechanism, does not have the effect one may expect: Manipulation rates were found to be surprisingly low even when the voter who had complete information about the other agent's preferences knew about his superior position. Only when the agent with superior information was also informed about the other agent's actual vote, manipulation rates went up significantly. This suggests that behavior in their context is affected by uncertainty, while distributional concerns, in particular inequality aversion, do not seem to play an important role. The authors conclude that the fear of strategic manipulation is not always justified in an applied framework.

To investigate the effect from the lack of collusion-proofness, we consider a simple setting in which all agents are informed about others' preferences in order to facilitate the occurrence of collusion. In addition, we will vary the possibility to communicate: In one treatment, agents make individual decisions without communication, and thus only tacit collusion is possible. In the other treatment, a particular communication network exists, which allows agents to communicate and co-ordinate their decisions within their subgroup. This setup shall give some insight into the question whether collusion can occur tacitly, i.e. without explicit communication, but simply from learning or observing others' behavior over time, or whether other motives, such as distributional concerns, prevent subjects from colluding when some subjects would suffer significantly from the collusive outcome.

In the literature on collusion in auctions it is a well-established fact that communication works as a coordination device (see e.g. the early experiments by Isaac and Walker 1985). Tacit collusion, on the other hand, seems to be much more difficult to establish, and even if it is observed, it is often unstable (see e.g. Burns 1985, or Isaac and Smith 1985 in an industrial organization context). More recently, Sherstyuk (1999) designed auction institutions with the particular rule that bidders are allowed to match the highest outstanding bid, and in case of multiple highest bidders the lottery decides on the winner. With this rule, stable tacit collusion was observed. In a later multi-object auction experiment, Kwas-

nica and Sherstyuk (2007) found tacit collusion, where payoff-superior collusive outcomes required complex strategies such as signalling and coordination across market. The experimental study by Burtraw et al. (2008) shows how institutional settings affect the likelihood of collusion in different auction types for emission permits; clock auctions are most susceptible to collusion with and without communication, as they allow bidders to focus on a single dimension. In the field of industrial organization, Haan et al. (2009) pointed out in a survey paper that without communication, firms have little success in establishing collusion, while the possibility to communicate yielded collusive outcomes. Overall, it seems that tacit collusion can be established only under certain conditions, and it is not clear how to generally describe such conditions. The experiments by Li and Plott (2009) find sufficient conditions under which tacit collusion among more than five players develops in auctions, and they use a repeated structure, which allows for instruments to sustain collusion, such as punishment through overbidding.

In contrast, our setup does not provide such additional instruments to promote collusion other than the complete information about preferences. We investigate the occurrence of tacit collusion by looking into a different mechanism where there are incentives for collusion, but no direct individual action towards a non-colluder is possible. We find little evidence for tacit collusion despite the existence of a simple collusive strategy. Our observations may be explained if subjects have varying beliefs over the choices of others and best-respond to a probability distribution of others' choices. In a behavioral model using quantal response equilibrium, we estimate a mixture of two possible error structures: subjects either believe that errors in choices are independent, which would predict more probability on the dominant strategies, or that errors of same-type subjects are correlated, which would predict the collusive outcome. This model helps to explain the absence of tacit collusion, and communication can be interpreted as a device to correlate individual beliefs.

## 2 Experimental Design

Consider a group of five members, who all have given preferences over four different alternatives Alpha (A), Beta (B), Gamma (C), and Delta (D). Each alternative  $a$  represents a non-excludable public good, and agents face the decision whether or not each public good should be provided. The group consists of two subgroups: the “majority” of 3 voters, referred to as M-type agents (indexed as  $m = M_1, M_2, M_3$ ), and the “minority” of 2 voters, referred to as N-type agents (indexed as  $n = N_1, N_2$ ). All members within a subgroup have identical (induced) preferences, but preferences across the two subgroups differ. The true valuation  $v_i^a$  of agent  $i$  for each alternative  $a$  is represented in Table 1. All agents have full information about

the valuations of the other agents and the structure of the group.

Table 1: Net valuations for the four alternatives

agent	Alpha	Beta	Gamma	Delta
M-type	30	-20	-30	10
N-type	10	40	-20	-40
group	110	20	-130	-50

The outcome is determined as follows: Each agent  $i$  submits a report  $r_i^a$  for each alternative  $a$ , which may or may not correspond to this agent's true valuation  $v_i^a$ . We shall allow for reported values between -60 and 60 in steps of 10.<sup>4</sup> An alternative  $a$  is selected (the public good is implemented) if and only if the sum of all reported valuations for  $a$  is strictly positive ( $\sum_i r_i^a > 0$ ). The number of selected alternatives is thus endogenously determined. Furthermore, if an agent is pivotal in the decision, he has to pay the Clarke tax, which corresponds to the total amount reported by all other agents, i.e. it reflects the cost that this agent imposes on the rest of society by changing the decision. Let  $r_{-i}^a$  be the vector of reports for alternative  $a$  omitting the report of agent  $i$ . Then the Clarke tax  $t_i^a$ , which agent  $i$  has to pay for alternative  $a$ , is calculated as follows:

$$t_i^a = \begin{cases} \sum_{j \neq i} |r_j^a| & \text{if } \sum_i r_i^a > 0 \text{ and } \sum_{j \neq i} r_j^a \leq 0, \text{ or} \\ & \text{if } \sum_i r_i^a \leq 0 \text{ and } \sum_{j \neq i} r_j^a > 0 \\ 0 & \text{otherwise} \end{cases}$$

Each subject in the experiment thus had to submit four reports, one for each alternative. The simultaneous reports then determined which of the four alternatives would be selected. The total payoff of a player is the sum of the payoffs from the selected (winner) alternatives minus the respective tax he has to pay. Subjects were explained on the instruction sheets how payoffs are calculated. In addition, we prepared detailed payoff tables that include the Clarke tax (see Figures 9-12 in the Appendix), so that subjects could also read off the tables what their payoff for each possible combination of chosen values would be.<sup>5</sup>

<sup>4</sup>The steps of 10 may render the reports space rather sparse, however, the dimensions of our payoff tables were already quite large with 13 columns and 49 rows. Extending it by only one more step in each direction would have immediately added 16 additional cells.

<sup>5</sup>From the payoff tables one can see that agents are indifferent between stating their true valuation and the next higher valuation for all alternatives, e.g. for M-types, the same payoff is reached for Alpha by reporting 30 or 40, for Beta with -20 and -10, etc. This is simply an artifact of the discrete space of reports, which implies that each agent has two weakly dominant strategies in the pivotal mechanism. For the choice of the socially efficient outcome it does not matter which of these two strategies is selected.

The valuations for each type are chosen such that there should be consensus regarding the provision of alternatives Alpha (positive valuation for both types) and Gamma (negative valuation for both types). A major conflict of interest between the two subgroups is expected for alternatives Beta and Delta. Since our research question regards susceptibility of the pivotal mechanism to collusion, we used a complete information setup that enhances the occurrence of collusion and gives simple and precise theoretical predictions. All subjects played 10 rounds of this game, and they remained within the same group for all rounds. They were paid out 5 cents for each point earned in the experiment. Since N-types were likely to make losses, they received a bonus of 15 Euros at the beginning of the experiment; losses were then subtracted from this bonus. Subjects were not informed at any time about the actual reports of any other subject. Feedback only included the total sum of reported values for each alternative and own payoffs from all four alternatives. We used communication as a control variable to differentiate between tacit and explicit collusion. In the *No Communication* treatment, subjects had no possibility of communication, while in the *Communication* treatment we allowed for communication within a subgroup but not between subgroups. This is sufficient in order for subgroups with identical preferences to coordinate their reports.<sup>6</sup> Communication was possible by using a chat program via computers, which closed after five minutes. Chatting was anonymous, and any sort of agreement made via chat is non-binding. The experiment was run on computers using the software *z-tree* (Fischbacher 2007) as well as the recruitment software by Greiner (2004). Average earnings were 10.30 Euros, and the duration of a session was about 45 minutes. A total of 80 subjects participated in this experiment at the University of Innsbruck, they were equally distributed between the two treatments.

### 3 Predictions

Agents in our game have to submit one report for each of the four alternatives. The total payoff of a player is the sum of the payoffs resulting from the decisions on each alternative. As a theoretical benchmark, we assume additively separable utility functions, and we may thus consider the decision for each alternative separately.<sup>7</sup> The pivotal mechanism is designed such that individual agents have an incentive to report their true valuation for an alternative. In the absence of collusion possibilities, with a discrete strategy space and a binary decision regarding the selection of an alternative each agent has two weakly dominant strategies for each alternative: reporting the true valuation or its next-highest level (see footnote

---

<sup>6</sup>Agents of different type have no incentive to collude by virtue of their opposed preferences for alternatives Beta and Delta.

<sup>7</sup>This shall serve as main benchmark. In the section describing the experimental results, we will also discuss the case where agents bundle the alternatives.

4). Reporting a dominant strategy would then ensure that the social optimum is achieved. In our experiment this means that only Alpha and Beta should be selected, since only for these two alternatives the sum of the valuations for all group members is strictly positive.

In games where collusion is beneficial for a subset of players, coordination of individual behavior is crucial to achieve the collusive outcome. The concept of strong equilibrium by Aumann (1959) requires that such a collusive agreement is not subject to an improving deviation by any coalition of players. A deviation is self-enforcing when there are no further profitable deviations for a subset of players. Aumann's (1959) strong equilibrium does not require deviations to be self-enforcing, i.e. an agreement has to be resistant to any deviation which itself is not required to be resistant to further deviations. This is a strong requirement for non-cooperative games like ours where pre-play communication is allowed, but agreements on coordinated actions are non-binding. Bernheim, Peleg and Whinston (1987) suggested that such agreements should be self-enforcing. Thus, they require for their notion of coalition-proof Nash equilibrium (CPNE) that an agreement is a Nash equilibrium and immune to improving deviations which are self-enforcing. We will use this equilibrium notion as a theoretical benchmark for the collusion case. Furthermore, we will focus on pure strategies as a theoretical benchmark for our experiment, as equilibria using mixed strategies require even more sophistication by players.

As Peleg (1998) showed for a two-person example, the profile of true preferences in the pivotal mechanism is not a CPNE, as other Nash equilibrium profiles exist that Pareto dominate truth-telling. In our setup, note that Alpha and Gamma do not impose any conflicts of interest upon the two types, since Alpha offers a positive and Gamma a negative payoff for all. Thus, in a CPNE we must have that Alpha is selected and no player has to pay the Clarke tax, while Gamma is not selected and no player pays a tax. If one or more players would have to pay a tax, they would prefer to revise their report such that they do not need to pay the tax. Such a deviation would be self-enforcing, as there exist profiles of reports for Alpha and Gamma such that no player has to pay a tax. For Alpha, a CPNE is thus a profile of reports such that  $\sum_i r_i^A - \max_i \{r_i^A\} > 0$ , while for Gamma, a CPNE is a profile of reports such that  $\sum_i r_i^C - \min_i \{r_i^C\} \leq 0$ . There is a large set of CPNE for Alpha and Gamma, but the *CPNE outcome* is unique. Due to the lack of conflict, these two alternatives are less interesting and shall serve mostly as a reference regarding subjects' understanding of simple decisions compared to the more complex ones for Beta and Delta.

Regarding the CPNE for alternatives Beta and Delta, note that by coordinating behavior the majority can ensure that their preferred outcome is implemented. For Beta, a symmetric strategy for all M-types of reporting  $-60$  ("maximally underreporting") leads to a CPNE outcome, since, irrespective of the N-types' reports, Beta is not selected and no M-type pays a tax, thus no coalition of M-



types can improve upon this outcome. A similar reasoning applies to Delta when the coalition of M-types maximally overreports. However, there are many other CPNE, which shall be characterized in the following.

- (i) *In a CPNE Beta must not be selected and Delta must be selected.* Suppose otherwise, i.e. suppose Beta is selected (Delta is not selected). This implies a payoff of  $-20$  ( $0$ ) for M-types. But the coalition of all M-types can guarantee a payoff of  $0$  ( $10$ ) for each M-type by maximally underreporting (maximally overreporting).
- (ii) *In a CPNE only coalitions consisting of agents of the same type need to be considered.* To see this, consider a profile of reports that induces a given outcome regarding Beta. Suppose now that there exists a coalition including an M-type and an N-type who deviate in order to improve upon this outcome. If Beta is selected after the deviation, this deviation cannot be self-enforcing, since the coalition of all M-types can always ensure their preferred outcome in which Beta is not selected and no M-type pays a tax. If Beta is not selected after the deviation, then an N-type would only participate in this coalition if he would have to pay a tax of more than  $40$  before. This means that the original outcome must have selected Beta (otherwise the N-type would not pay a tax). But then it is again sufficient to consider only the coalition of all three M-types who can ensure to avoid the selection of Beta. A similar argument applies to Delta.
- (iii) *In a CPNE no N-type pays a tax.* Since N-types do not get their preferred outcome with regard to Beta and Delta in a CPNE, they cannot be pivotal since that would mean that they submitted a report that goes against their preferred outcome.
- (iv) *A CPNE in which all three M-types pay a tax does not exist.* Suppose otherwise, then the coalition of all M-types could improve by maximally underreporting for Beta and maximally overreporting for Delta, in which case no M-types pay a tax.
- (v) *In a CPNE we must have for Beta  $\sum_m r_m^B \leq -100$  and for Delta  $\sum_m r_m^D \geq 100$ , for  $m = M_1, M_2, M_3$ .* Recall that in a CPNE Beta must not be implemented. To achieve this, it is sufficient for M-types to submit reports such that the tax for N-types would be higher than the benefit of having Beta selected. When  $\sum_m r_m^B \leq -100$ , at least one N-type would have to pay a tax of more than  $40$  and thus prefer not to have Beta implemented. A similar argument applies to Delta.

In a CPNE, M-types thus submit reports such that either one, or two, or none of the M-types pay a tax. Furthermore, if two M-types pay a tax, at least one

of them must report -60. Otherwise they could decrease their tax by reporting -60 and would thus both be strictly better off. This characterization of CPNE for Beta and Delta helps to identify all CPNE of the one-shot game (by computer programming); there are 26195 CPNE for Beta and 25484 CPNE for Delta.

Thus, leaving aside mixed strategy equilibria, we already have a large number of pure-strategy CPNE. Furthermore, our experimental design may induce subjects to consider the 10 repetitions of the described one-shot game as one repeated game, since subjects remain within a given group for all 10 periods. If we consider this as a repeated game, the number of equilibria becomes even larger. This is a drawback of our design, however, we thought that it was important to give subjects the possibility to learn from past experience, and since learning here involves coordinating beliefs about others' choices, this is most easily achieved within a constant group. Otherwise, with changing partners, behavior of others may simply become too difficult to predict. Our focus in this experiment was to study the occurrence of tacit collusion, and knowing that this is difficult to achieve in most contexts (see discussion in Introduction), we wanted to facilitate the conditions for collusion and allow for learning within a group, so that we also can consider decisions of more experienced players.

Clearly, coordination on any one equilibrium is very difficult, as players must make their choices based on their beliefs about other players' choices. However, amongst the large set of CPNE we can find one that seems particularly appealing: the CPNE with symmetric reports on the boundary of the strategy space for each M-type, i.e. when players maximally underreport for Beta and maximally overreport for Delta. These extreme reports avoid tax payments and ensure the preferred outcome for the coalition of M-types *independent* of the reports of N-types. Therefore, they are weakly dominant for the coalition of M-types and should be of particular appeal in an experimental setup. Playing this CPNE yields a payoff of 40 to M-types, while if all played their individually dominant strategies would yield a payoff of 10 to M-types. We are interested in observing whether collusion can be established, and this seems a particularly simple way to achieve such collusive outcome, as it avoids difficulties in coordination. Furthermore, we know from experimental evidence in Charness et al. (2007) that salient group membership induces individual behavior towards more favorable outcomes for other group members, therefore, by the design of two subgroups with partly opposed interests, belonging to the majority would imply that M-types prefer the symmetric extreme reports to any other CPNE profile where some agent risks to pay a tax. And if our design of majority versus minority does create a feeling of competition towards the other subgroup, it would also lead us to expect better coordination, as Bornstein et al. (2002) have shown in their intergroup competition in a coordination game. While one may expect that this CPNE with extreme reports for M-types can also occur in the treatment without communication due to its simplicity and symmetry,

it seems rather unlikely that any other (asymmetric) CPNE strategy profile can be achieved without explicit communication. Therefore, we take CPNE with symmetric reports for the M-types as a benchmark for identification of possible tacit collusion in the treatment without communication. Since any report of N-types is a best response to this weakly dominant strategy for the coalition of M-types, we have  $13^2 = 169$  of these equilibria for both Beta and Delta, in which the strategies of all M-types are identical.

## 4 Experimental Results

### 4.1 Outcomes in the NoCommunication-Treatment

Group	Round									
	1	2	3	4	5	6	7	8	9	10
1	AB	ABD	ABD	ABD	ABD	AD	ABD	AB	ABD	ABD
2	AD	AD	AD	AD	AD	ABD	AD	AD	AD	AD
3	ABD	AD	ABD	AD	AD	AD	ABD	AB	A	ABD
4	AD	ABD	ABD	ABD	ACD	ABD	AB	ABC	AD	AD
5	AB	AB	AB	AB	AB	ABD	ABD	AB	AB	AB
6	AB	ABD	AB	AB	AB	ABCD	AD	AB	AB	AB
7	AD	ABD	AD	ABD	AB	AB	AD	ABD	AB	AB
8	AD	ABD	ABD	ABCD	ABD	AB	AB	AB	AB	AB

(A=Alpha, B=Beta, C=Gamma, D=Delta)

Figure 1: No Communication: All Outcomes over time

**Selected alternatives.** Recall that an alternative  $a$  is selected if the sum of all reports is strictly positive:  $\sum_i r_i^a > 0$ . Figure 1 shows the various combinations of selected alternatives in each group for each of the 10 rounds in the treatment without communication. First, notice that Alpha, which is the only alternative that yields a positive payoff for both types, was always selected. Gamma, the only alternative that yields a negative payoff for both types, was selected in less than 5%, which is in the range of typical errors of subjects in experiments. Regarding the two alternatives where preferences are opposite, Delta, the majority's preferred outcome, was selected in 63%, while Beta, the minority's preferred outcome, was selected in 70% of all games. Interestingly, Beta and Delta were selected simultaneously in a significant number of games. It is easily seen that three combinations of outcomes are predominant in this treatment: the social optimum {Alpha, Beta} occurred in 35%, the best possible collusive outcome for M-types {Alpha, Delta} occurred in 27.5%, and then we observe the combination {Alpha, Beta, Delta} in 31% of all outcomes in this treatment. The first important results are thus that

(i) tacit collusion does not easily emerge when communication between agents is not possible, despite full information and a seemingly easy-to-achieve, symmetric collusive outcome, and (ii) the social optimum seems to play an important role.<sup>8</sup>

The third prominent outcome, {Alpha, Beta, Delta}, requires some attention, since, at first sight, it is not clear whether it occurred due to unsuccessful attempts of tacit collusion or whether subjects possibly considered it a desirable outcome. The latter could be the case for subjects with distributional concerns such as a self-centered inequality aversion as introduced by Fehr and Schmidt (1999), and further extended by Charness and Rabin (2002). If the triple {Alpha, Beta, Delta} is regarded as a bundle, this bundle offers a total payoff of 20 for M-types and 10 for N-types, thus being the bundle with the least unequal positive payoff for the two types of players. In particular, this bundle still favors the majority and may thus be more easily supported by M-types. Compared to this bundle, the collusive outcome {Alpha, Delta} implies a negative payoff of -30 for the minority; the social optimum {Alpha, Beta} would imply positive payoffs for both types, but they are highly unequal and favor the minority. The bundle {Alpha, Beta, Delta} implements payoffs that are as close as possible to equal positive payoffs for all agents. If there are subjects with distributional preferences, this may well represent their preferred choice.<sup>9</sup> To understand whether this outcome, which occurred in over 30% in the treatment without communication, came about as a purposeful choice or if it happened due to insufficient coordination, we look at outcomes over time.

**Outcomes over time.** Figure 1 offers a first impression of how possible learning through repeated interaction affects outcomes. We define the first three rounds as “early” rounds and the last three as “late” rounds in the experiment. Then the outcome {Alpha, Beta, Delta} occurs significantly more often in earlier rounds than in later rounds (Pearson  $\chi^2$ :  $p = 0.06$ ). This would point towards an unintended outcome, since experience leads away from selecting this bundle. In other words, inequality averse preferences do not play an important role in this setup, as the only choice which would be consistent with such preferences loses importance over time. This observation is in line with the findings of Puppe and Kube (2009) for the Borda count method. Our results below for the treatment with communication further confirms this result, as this bundle plays no role in the observed outcomes.

A comparison of the frequency of selected outcomes over time further confirms

---

<sup>8</sup>In the section on behavior below, we will discuss whether {Alpha, Beta} is a frequent outcome because it is socially optimal or whether this is due to behavior that arises from other incentives here.

<sup>9</sup>For example, in the Fehr-Schmidt (1999) model of inequality aversion this would be the preferred choice of an N-type agent if the parameter  $\beta$ , which measures the weight the agent puts on others’ monetary payoffs when he is behind, is sufficiently high.

that M-types did not succeed in establishing tacit collusion on {Alpha, Delta}. There is no significant difference in the occurrence of {Alpha, Delta} in early and late rounds (Pearson  $\chi^2$ :  $p = 0.33$ ), i.e. collusion does not become more prevalent even though subjects gain experience and learn about other agents' reporting behavior in their own group. Learning or experience alone was thus not sufficient to establish collusion amongst M-types. On the other hand, the outcome {Alpha, Beta} gains importance with subjects' experience: it is selected significantly more often in later rounds compared to earlier rounds (Pearson  $\chi^2$ :  $p < 0.05$ ). In late rounds, {Alpha, Beta} was selected in 54% of all group decisions, while in early rounds it was selected only in 25%. This proportion corresponds to the findings of implementation of the first-best outcome through the pivotal mechanism in other experiments where collusion was not possible (e.g. Cason et al. 2006). Achieving social efficiency might, however, not be due to the pivotal mechanism's incentive to submit truthful reports (or at least to select a dominant strategy), but rather due to the already noted absence of tacit collusion for M-types, while N-types are more successful in this regard and {Alpha, Beta} is their preferred outcome. We will refer to this point in the section on subjects' behavior below.

		all subjects			tax payers		
		mean profit	Std.Dev.	# obs	mean profit	Std.Dev.	# obs (frac)
Beta	M-type	-17.46	9.67	240	-24.59	14.45	37 (.15)
	N-type	22.43	19.28	160	11.29	14.32	31 (.17)
Delta	M-type	3.50	8.50	240	-10.00	12.75	33 (.14)
	N-type	-28.56	17.26	160	-25.42	13.18	24 (.13)

Table 2: No Communication: Profits by Type

**Payoffs and Tax payments.** Describing outcomes only in terms of selected alternatives is not sufficient in order to evaluate the performance of the pivotal mechanism and the occurrence of collusion. Since tax payments may be considerable, overall payoffs are also relevant. Furthermore, payoffs over time should give information about subjects' possible adjustments after observing outcomes and payoffs in previous rounds. For Alpha and Gamma, there were overall few tax payments (less than 5%). Since Beta was selected in over 70%, it is important to understand at which cost for N-types this outcome came about. Table 2 shows that of 160 decisions of N-types, a tax was paid in 17% to have Beta implemented. However, the tax-paying subjects are significantly better off despite the imposed Clarke tax (Sign test:  $p < 0.01$ ; Wilcoxon signed rank (WSR):  $p < 0.01$  for averages by individual) compared to the case where Beta is not selected. The same is true for N-types who paid a tax in order to avoid Delta. Thus, while one would normally expect that subjects learn to avoid tax payments over time since they decrease payoffs, the feedback here gives tax-paying N-types a positive reinforcement

to overreport for Beta and underreport for Delta. It is then not surprising that the frequency of tax payments for Beta does not change over time. For M-types, the picture is different: Subjects who paid a tax in order to avoid Beta are no better off than in the case where Beta is selected (Sign test:  $p = 0.38$ ). For Delta it is even worse: M-type subjects who paid the Clarke tax to have Delta selected are significantly worse off than if it had not been selected (WSR and Sign test:  $p < 0.01$ ). Yet, even though it was never profitable for M-types to pay a tax, the frequency of tax payments over time does not change significantly. Thus, on the one hand, individual subjects in the role of M-types seem to keep trying to overreport - or underreport, on the other hand, M-types altogether thus do not seem to learn to avoid tax payments, even though as majority they could have done so with a simple coordination on extreme reports. It is precisely this coordination that does not improve over time.

## 4.2 Outcomes in the Communication Treatment

		Round									
Group		1	2	3	4	5	6	7	8	9	10
1	A	AB	AD	AD	AD	AD	AD	AD	AD	AD	AD
2	AB	AD	AD	AD	AD	AD	AD	AD	AD	AD	AD
3	AD	AD	AD	AD	AD	AD	AD	AD	AD	AD	AD
4	AD	AD	AD	AD	AD	AD	AD	AD	AD	AD	AD
5	AB	AD	AD	AD	AD	AD	AD	AD	AD	AD	AD
6	ABD	AC	AB	AB	AB	AD	AD	ACD	AD	AD	AD
7	ABD	A	AD	AD	A	AD	AD	AD	AD	AD	ABD
8	AD	AD	AD	AD	A	AD	AD	AD	AD	AD	AD

(A=Alpha, B=Beta, C=Gamma, D=Delta)

Figure 2: With Communication: All Outcomes over time

**Selected alternatives.** When communication among same-type subjects is introduced, outcomes change dramatically, and the pivotal mechanism breaks down completely. Figure 2 shows that the majority's payoff-maximizing outcome {Alpha, Delta} occurs in 82.5% of all games, while other outcomes, in particular the socially efficient {Alpha, Beta} plays no role. While tacit collusion did not seem to work for the majority when communication was not possible, collusion is now widely established. Over time, the occurrence of the majority's payoff-maximizing outcome {Alpha, Delta} increases significantly from 63% in early rounds to 92% in late rounds (Pearson  $\chi^2$ :  $p < 0.02$ ). Since the social optimum is now rarely observed as an outcome, this means that social efficiency was most likely not a goal *per se* in the treatment without communication; outcome {Alpha, Beta} rather occurred there because the minority was able to implement it as their preferred outcome.

Regarding bundle {Alpha, Beta, Delta}, which was frequently selected in the treatment without communication, we can finally conclude that it must have occurred unintentionally, i.e. due to insufficient coordination of M-types. It disappeared completely in this treatment where subjects could have explicitly decided to vote for it. This leads us to the conclusion that social preferences, which could have influenced outcomes, have no importance in this context. One may argue that allowing for communication only among subgroups may have emphasized competition between groups and selfish behavior. This is possible, and it might be interesting to investigate this effect in further experiments. Our main goal, however, was to understand whether collusion is easily established in a pivotal mechanism context, and we thus chose a design with two treatments on the extremes regarding communication possibilities.

**Payoffs and Tax payments.** Comparing the two treatments with regard to tax payments, Table 3 shows that overall fewer agents paid taxes in the treatment with communication. Communication thus improved the efficiency regarding tax payments: for M-types in the Communication-treatment, there were significantly less tax payments for both Beta (Pearson  $\chi^2$ :  $p = 0.02$ ) and Delta (Pearson  $\chi^2$ :  $p < 0.01$ ). For N-types, the number of subjects who paid a tax for Beta also decreased significantly in the Communication-treatment (Pearson  $\chi^2$ :  $p < .01$ ), only for Delta the decrease is not significant (Pearson  $\chi^2$ :  $p = 0.12$ ). In accordance with Figure 2, mean profits of M-types and N-Types also changed dramatically, since now the majority was mostly able to implement their preferred outcomes.

		all subjects			tax payers		
		mean profit	Std.Dev.	# obs	mean profit	Std.Dev.	# obs (frac)
Beta	M-type	-4.88	11.05	240	-30.00	16.12	21 (.09)
	N-type	1.94	8.58	160	5.83	14.43	12 (.08)
Delta	M-type	7.96	5.05	240	-9.00	7.38	10 (.04)
	N-type	-38.31	8.18	160	-35.33	14.07	24 (.09)

Table 3: Communication: Profits by Type

Overall, the observations so far show that outcome {Alpha, Delta}, which yields the highest monetary payoff for the majority, is chosen in the treatment with communication even though it yields a highly unfavorable outcome for the minority. Since communication would have allowed to coordinate on any other outcome (if so desired), we must conclude that {Alpha, Delta} must have also been desirable in the NoCommunication treatment, but was not attained due to the lack of communication. Communication thus seems to work as a coordination device in the Pivotal Mechanism, a result that is well-established in the literature on collusion in auctions. We also observe that coordination among two agents (N-types) works better in the NoCommunication treatment than coordination among

three agents (M-types). However, looking more carefully into why this is the case, the section below on behavior will show that this may also be due to the expected losses M-type agents face. The main question which remains to be answered is why subjects in our Pivotal mechanism experiment did not succeed in establishing tacit collusion. For this purpose, we take a closer look into subjects' individual behavior.

### 4.3 Individual Behavior

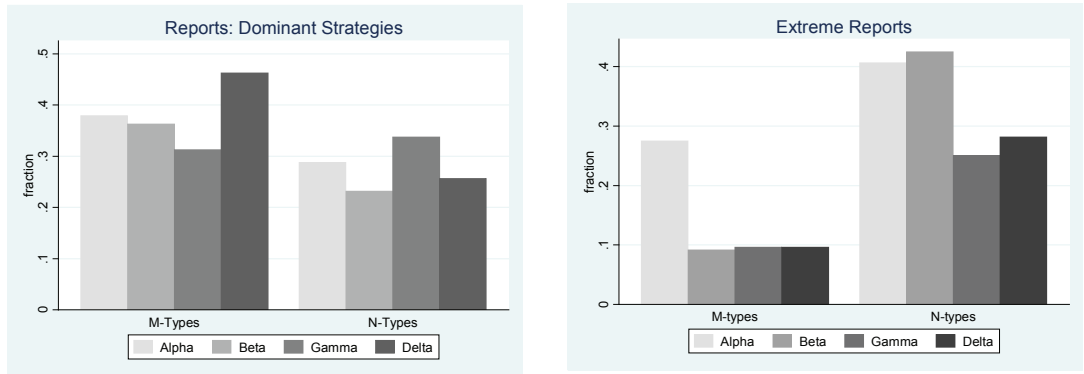


Figure 3: Rel. Frequencies of Dominant Strategies and Extreme Reports (NoCom)

**Descriptive Behavior.** Analyzing individual reports should reveal why tacit collusion by M-types on the outcome  $\{\text{Alpha}, \text{Delta}\}$  was not observed in the No-Communication treatment even though the Communication treatment showed that it was clearly a desirable outcome for M-types. Individual choices also will reveal to which extent the pivotal mechanism was successful in eliciting truthful behavior. The left graph of Figure 3 shows the fraction of dominant strategies for each alternative by player type in the NoCommunication treatment.<sup>10</sup> For M-types, the fraction of reports that are consistent with a dominant strategy of the non-collusive game is between one third and one half for all alternatives. Notice, in particular, that the dominant strategy choices for Beta and Delta, where more strategic reports might be expected, are not lower than those for Alpha and Gamma. For N-types, the proportion of dominant strategy reports is overall lower. In particular, we found significant differences between the fraction of dominant strategy reports of the two types for all alternatives except Gamma (MWU for average choices by individuals:  $p = 0.06$  for Beta,  $p = 0.02$  for Delta). For subjects who do not use dominant strategies, the question is then whether they report strategically

<sup>10</sup>Recall that the pivotal mechanism provides two weakly dominant strategies for an individual in this experiment if we consider only non-collusive behavior: reporting the true valuation and reporting its next-highest value.



instead. The right graph of Figure 3 displays the proportion of extreme reports, i.e. the proportion of subjects that reported the highest possible value for an alternative with a positive valuation and the lowest possible value for an alternative with a negative valuation. While almost 30% of M-types maximally overreport for Alpha, where strategic behavior does not seem necessary as there is consensus over this alternative, the proportions of extreme reports for Beta and Delta are rather low. N-types, on the other hand, submit extreme reports more frequently (MWU for differences between the two types:  $p < 0.01$  for both Beta and Delta for average choices by individuals). At first sight, this is surprising, since extreme reports of all three M-types would ensure a favorable outcome of the corresponding alternative at no tax, while in the case of N-types there is no profile of reports that leads to a certain outcome. Thus, on the one hand, M-types are in the better position, since being the majority they can enforce their preferred outcome if they all collude. On the other hand, it might just be easier to coordinate two N-type reports as opposed to three M-types when collusion can only occur tacitly.<sup>11</sup>

There is a further argument that may explain the reluctance of M-types to submit extreme reports in this treatment. As already noted in the predictions, there is a large number of equilibria, and subjects' beliefs about the actions of the other agents in their group become important. If we now consider that an agent is best responding to a probability distribution over possible actions of the other four group members, the payoff tables (see Appendix) show the following regarding expected losses from extreme reports: (i) For M-types, there are 4 possible combinations of own and others' reports (cells in payoff table) for which an extreme report is worse than a dominant strategy report for both Beta and Delta. (ii) For N-types, there is only one possible combination of reports for Beta in which an extreme report gives a lower payoff than a dominant strategy report, while for Delta there are two such possibilities. In addition, the difference in losses is larger for M-types. Thus, the expected losses from extreme reports differ for the two types, and one may therefore expect that N-types are more likely to submit extreme reports due to their smaller expected losses. Note that this is due to our parametrization, i.e. if the true valuations for N-types were more in the interior of the reports space, there would be more states in which losses can occur for N-types, and this may have affected their incentives for extreme reports. Our specific parametrization thus created more favorable conditions for N-types to collude, which may be of behavioral relevance, even though it plays no role for the concept of CPNE since, in theory, it does not alter the majority's possibility to collude. For the robustness of our results we shall therefore keep in mind that the behavioral path towards

---

<sup>11</sup>This observation finds some support in the literature: Haan et al. (2009) found that tacit collusion is found to some extent for industries with only two firms. In the experiment of Kwasnica and Sherstyuk (2003), collusion occurs only in markets with two bidders; with five bidders no collusion was observed. Experience increased the incidence of collusion in the 2-bidder-market, while it did not help when experience was gained in the 5-bidder-market.

collusion for M-types might have been affected by this advantage for N-types.

**Behavior over time.** Since coordinating reports might require some rounds of learning, we check whether the rate of extreme reports changes over time. Table 4 shows that in the treatment without communication, there is no significant difference in extreme reports for M-types between early and late rounds; note that the fraction of M-types who report extreme values in NoCommunication is overall very small. For N-types, extreme reports are somewhat higher in late rounds, but the difference is not significant. Overall, there are no significant learning effects for either type in the treatment without communication; allowing subjects to gain experience is not sufficient, in particular for M-types, to behave strategically despite the complete information context.

Table 4: Extreme reports over time

Treatm.	Altern.	M-types			N-types		
		early rounds	late rounds	all rounds	early rounds	late rounds	all rounds
NoCom	Beta	.083	.111	.092	.375	.479	.425
	Delta	.097	.111	.095	.250	.354	.281
Com	Beta	.556	.722	.642	.479	.500	.544
	Delta	.528	.861	.733	.479	.458	.456

In the treatment with communication, Table 4 shows that over one half of all M-types now report  $-60$  for Beta and  $60$  for Delta in early rounds, and this proportion significantly increases to  $72\%$  for Beta and  $86\%$  for Delta in late rounds (Pearson  $\chi^2$ :  $p < 0.03$  for Beta and  $p < 0.01$  for Delta). This learning effect implies that it is not only the lack of communication that prevents M-types to collude, but that probably also cognitive limits constraint subjects from choosing their collusive best response in this situation. We observe the particularity that subjects are more hesitant with extreme reports for an alternative with a negative valuation. N-types, on the other hand, do not use extreme reports more extensively than in the treatment without communication, and there is also no change over time for extreme reports of Beta and Delta. This is easily explained: as the collusive behavior of M-types now determines the outcome, the reports of N-types become irrelevant.

We gained some additional insight regarding selfishness and fairness concerns by analyzing the chat protocols. Only in 3 out of 8 subgroups of M-types unfairness towards N-types was mentioned by single subjects. Subjects who brought up the fairness topic did not insist much and the issue was discarded rather quickly. This is in line with the findings of Bosman et al. (2006), who analyzed video taped

discussions preceding group decision making in a power-to-take experiment. In their context, fairness is also discussed very little and fairness standards are prone to the self-serving bias. As hypothesized earlier, we can also see from the chat protocols that not all M-types subjects are able to understand strategic behavior here. This is why learning is significant only in the communication treatment, i.e. only with the explicit help of others can some subjects follow the collusive strategy. Communication thus helps subjects to understand a complex problem and it helps coordinate strategies towards a collusive outcome.

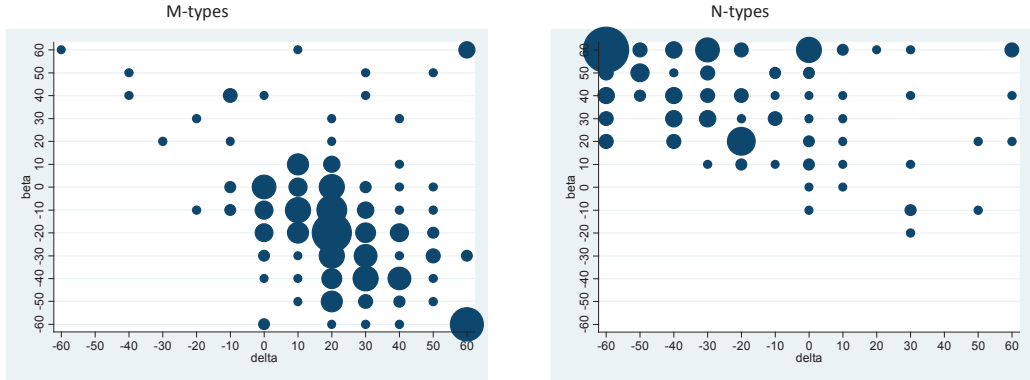


Figure 4: No Communication: Individual Reports for Beta and Delta

If we assume that individuals have a common approach for decision making for all alternatives, i.e. they are consistent in their choices regarding the criteria of decision making, we should consider a behavioral model that explains choices for all alternatives. Figure 4 plots the observed frequency of individual choices for each possible combination of reports for Beta and Delta (larger dots indicate a larger number of observations). On an individual level, the vast majority of subjects in the role of M-types choose reports that are positive for Delta and negative for Beta, but there is considerable dispersion of these individual choices. Many M-type subjects choose reports that correspond to one of the two weakly dominant strategies of the non-collusive game for Beta and/or Delta:<sup>12</sup> About one quarter of M-types' reports are consistent with a dominant strategy of the non-collusive game for both Beta and Delta. For N-types, the proportion of dominant strategy reports is considerably lower at 5%. The large dot on  $(-60, 60)$  represents the 20% of extreme N-types' reports for both Beta and Delta, while the corresponding proportion for M-types is 7%. This gives a first explanation as to why  $\{\text{Alpha}, \text{Beta}\}$  is a frequent outcome in the treatment without communication. There are too few attempts by M-types on the group level to establish tacit collusion, and

<sup>12</sup>Note that four dots in Figure 4 represent these choices in the Delta-Beta-space:  $(10, -10)$ ,  $(10, -20)$ ,  $(20, -10)$  and  $(20, -20)$ .

this helps N-types who report extreme values more often and thus manage to implement their preferred choice despite being in the minority. {Alpha, Beta} thus is not chosen because the Pivotal mechanism gives incentives to implement the social optimum, but it is rather a result of the minority’s collusive behavior and the majority’s non-collusive behavior.

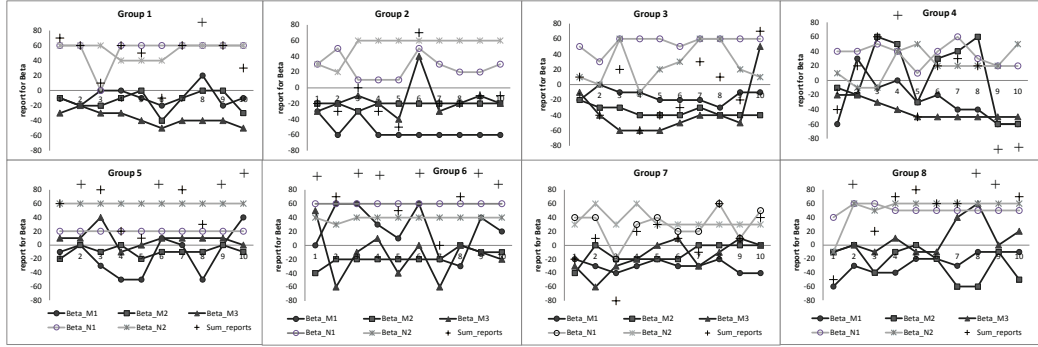


Figure 5: NoCommunication: Reports by group over time for Beta

For M-type subjects, the dispersion of individual reports for Beta and Delta indicates that M-types are far from reaching coordination on behavior here. The heterogeneity in strategies of M-types is thus the main reason for the various observed outcomes. Figures 5 and 6 show the reports individual subjects chose within their group over the 10 periods. The “+” sign marks the sum of all reports in the respective period. From Figure 5 one can see that in 5 out of 8 groups there is at least one of the two N-types who submits the extreme report for Beta in almost all periods, while for M-types this is not the case. In addition, there seems to be no convergence towards coordinated behavior over time for M-types. But given that we observe this heterogeneity in behavior which, on the aggregate, does not follow any of the predictions, we now consider a behavioral model that might support our data: quantal response equilibrium, originally introduced by McKelvey and Palfrey (1995) and further developed by Goeree et al. (2005).<sup>13</sup>

**Quantal Response Equilibrium (QRE).** In QRE, players’ best response functions are probabilistic, i.e. they choose among strategies based on their relative expected utilities, and the probabilities of choosing a strategy depends on the relative costs of the errors measured in expected payoffs. We use the Logit Equilibrium, the most commonly used version of QRE, where errors are assumed to follow a logistical distribution. Following the notation of McKelvey and Palfrey (1995), let  $s_{ij}$  denote pure strategy  $j$  of player  $i$ , where  $j = 1, \dots, 13$  and  $i = 1, \dots, 5$ , let  $\pi_{ij} = \pi_i(s_{ij})$  be the probability with which player  $i$  chooses pure

<sup>13</sup>We are grateful to an anonymous referee for suggesting this approach for the interpretation of our data.

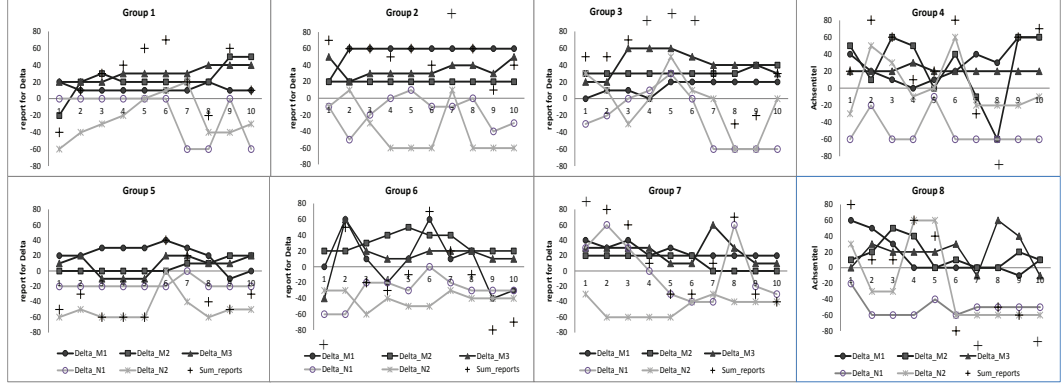


Figure 6: NoCommunication: Reports by group over time for Delta

strategy  $s_{ij}$ , and  $\pi = (\pi_1, \dots, \pi_5)$ , where  $\pi_i = (\pi_{i1}, \dots, \pi_{i13})$  is a mixed strategy of player  $i$ . When  $(s_{ij}, \pi_{-i})$  represents a profile where  $i$  chooses strategy  $s_{ij}$  and all other players choose their components of  $\pi$ , the corresponding expected payoff of player  $i$  is  $\bar{u}_{ij}(\pi) = u_i(s_{ij}, \pi_{-i})$ . Incorporating a player's random error, we then have  $\hat{u}_{ij}(\pi) = \bar{u}_{ij}(\pi) + \epsilon_{ij}$ . Players best-respond, i.e. they choose strategy  $j$  such that  $\hat{u}_{ij} \geq \hat{u}_{ik}, \forall k = 1, \dots, 13$ . Then the Logit Equilibrium requires that the probability that player  $i$  will use strategy  $j$  equals

$$\pi_{ij} = \frac{e^{\lambda \bar{u}_{ij}(\pi)}}{\sum_{k=1}^{13} e^{\lambda \bar{u}_{ik}(\pi)}}$$

The precision parameter  $\lambda$  is inversely related to the errors players make. In the following econometric model, we denote by  $I_P$  the set of individuals characterized by preferences of type  $P \in \{N, M\}$ . Given that we observe heterogeneity in subjects' behavior, e.g. some subjects who seem to follow a collusive strategy while others choose dominant strategies, we estimate a mixed model where we assume that each individual can be of two types: One type plays according to "non-collusive QRE", where the probability distributions on the actions are considered independent across players, while the other type plays according to "collusive QRE", where players of the same type play the same action. We assume that for each kind of preferences  $P \in \{N, M\}$  each type is drawn from a common prior distribution where the non-collusive QRE-type has probability  $\alpha_P$  and remains constant for all 10 periods. Let  $\pi_{ij}^{nc}(\lambda^{nc}) \equiv \pi_{ij}^{nc}$  be the probability that individual  $i$  chooses pure strategy  $s_{ij}$  in a non-collusive QRE with parameter  $\lambda^{nc}$ , and  $\pi_{ij}^c(\lambda^c) \equiv \pi_{ij}^c$  the probability that individual  $i$  chooses pure strategy  $s_{ij}$  in a collusive QRE with parameter  $\lambda^c$ . Let  $\hat{S}_i$  be the set of pure strategies that by player  $i$  actually chooses,  $\pi_i^{nc} = \prod_{s_{ij} \in \hat{S}_i} \pi_{ij}^{nc}$  be the probability of the observed

actions of player  $i$  of the QRE-type, and let  $\pi_i^c = \prod_{s_{ij} \in \bar{s}_i} \pi_{ij}^c$  be the probability of the observed actions of player  $i$  of the collusive QRE-type. The log likelihood on the whole sample is:

$$\ln L(\alpha, \lambda^i, \lambda^c | x) = \sum_P \sum_{i \in I_P} \ln(\alpha_P \cdot \pi_i^{nc} + (1 - \alpha_P) \cdot \pi_i^c), \quad \forall P \in \{N, M\}$$

We estimate this model for each alternative assuming that parameter  $\alpha_P$  can be different in the two treatments, i.e.  $\alpha_P = \gamma_{0,P} + \gamma_{1,P}d$ ,  $P \in \{M, N\}$ , where the dummy  $d = 1$  if and only if the treatment is with communication. For simplicity, we will write  $\alpha_P$  if  $d = 0$  and  $\alpha_P^{Chat}$  for  $d = 1$  in the estimations reported below.

	<b>Alpha:</b> Log Likelihood=-1603.00				<b>Beta:</b> Log Likelihood=-1538.86			
	Coef.	Std.Err.	z	$P >  z $	Coef.	Std.Err.	z	$P >  z $
$\lambda^{nc}$	.891	.282	3.15	.002	2.399	.337	7.11	.000
$\lambda^c$	4.194	.621	7.75	.000	2.660	.206	12.86	.000
$\alpha_M$	.219	.152	1.43	.152	.764	.079	9.55	.000
$\alpha_M^{Chat}$	.093	.059	1.56	.119	.428	.102	4.17	.000
$\alpha_N$	.171	.126	1.36	.173	.818	.114	7.16	.000
$\alpha_N^{Chat}$	.103	.093	1.11	.268	.769	.196	3.92	.000
	<b>Gamma:</b> Log Likelihood=-1766.61				<b>Delta:</b> Log Likelihood=-1557.39			
	Coef.	Std.Err.	z	$P >  z $	Coef.	Std.Err.	z	$P >  z $
$\lambda^{nc}$	1.041	.235	4.42	.000	2.244	.266	8.43	.000
$\lambda^c$	3.051	.383	7.96	.000	5.090	.356	14.27	.000
$\alpha_M$	.626	.140	4.47	.000	.801	.065	12.23	.000
$\alpha_M^{Chat}$	.214	.091	2.34	.019	.540	.102	5.29	.000
$\alpha_N$	.182	.117	1.56	.119	.660	.205	3.21	.001
$\alpha_N^{Chat}$	.365	.173	2.11	.035	.341	.185	1.84	.066

Table 5: Quantal Response Equilibrium Estimation for all alternatives

As can be seen from Table 5, this QRE-estimation that takes a mixture of a model with uncorrelated errors of individuals and a model with perfectly correlated errors of individuals of the same type reflects many of the observations we found in the data: The mixture parameter  $\alpha_M$  is high, in particular for Beta and Delta, which means that more observations of M-types' behavior in the treatment without communication correspond to a "regular" QRE with uncorrelated errors, where higher probabilities are placed on the weakly dominant strategies, which are in the interior of the strategy (reports) space. Only in case of alternatives Alpha and Gamma some estimates for the parameters of the econometric model are not significant, but note that these are the two alternatives where preferences of M-types and N-types were aligned, thus there was little expected punishment for behavior that deviated from best responses. Moving to the treatment with communication, the parameter  $\alpha_M$  decreases (difference is significant at  $p < .05$  for

all alternatives except Alpha), meaning that now the collusive QRE better fits the data, which is precisely what we noted above in the descriptive part on behavior. In collusive QRE, the estimates from our model would then imply that M-types place a probability of 55% on the extreme values for Beta and 65% for Delta. Communication thus serves to improve the correlation of strategies for M-types. For N-types, note that both QRE and collusive QRE predict more probability mass on the extremes of the strategy space (on the positive side for Beta and on the negative for Delta); in case of the QRE this is because of the already discussed reason that the dominant strategy lies close to the extreme reports for both alternative; this explains why the estimated parameter  $\alpha_N$  is relatively high when we expect that collusive QRE (and thus a low  $\alpha_N$ ) would better fit the data, and the differences to  $\alpha_N^{Chat}$  are not significant. We also observe that for N-types, QRE predicts a higher probability mass for more extreme values of Beta than Delta: 35% of all reports are estimated to be higher than the truthful report of 40 for Beta, while only 27% are lower than the truthful report of  $-40$  for Delta. This is due to the already mentioned fact that there are two weakly dominant strategies for each alternative: the truthful report and the next higher report. In the case of Delta, these are  $-40$  and  $-30$ , which shifts more probability mass towards the interior rather than towards the extreme.

Overall, using QRE as a behavioral model explains many of the facts we observe in subjects' behavior, in particular, it helps explain why M-types do not manage to have their preferred choices implemented. Assuming that subjects best respond to some probability distribution, the model with uncorrelated errors predicts more choices in the interior of the strategy space rather than at the corresponding extremes for M-types. Since subjects observed only the total sum of reports for each alternative, specific signals to same-type subjects were not possible, as they could not be read as such. Thus, repetition did not improve tacit collusion. If we assume that the QRE model above provides a good description of observed behavior, then we can conclude that best responding to a probability distribution (or, alternatively, assuming errors in expected payoff calculations) in our experiment does not support tacit collusion for the majority, since such deviations from the deterministic best response (in terms of CPNE) imply that several states of losses have to be taken into account in the expected payoff function used in the QRE. Then it is better for M-types to place a higher probability on the (individually) dominant strategies if correlation of strategies is not explicitly possible. If the payoff structure had been such that M-types had their true valuations been more at the extremes of the admissible reports space, we could have observed collusion among M-types. But that also would have been an artifact of the parametrization. In our experimental setup, we can conclude that without facilitating devices and no means to signal or punish a certain behavior, complete information about all players' preferences is not a sufficient condition for collusion of a majority.

## 5 Conclusion

This paper studies the susceptibility of the pivotal mechanism with respect to manipulation by groups. Knowing that this mechanism is not collusion-proof, it seems important to understand under which circumstances this property is responsible for the failure of the mechanism in implementing the social optimum. In a lab experiment where a group decides on the implementation of various alternatives, we investigate the occurrence of tacit and explicit collusion by allowing for communication in one treatment and prohibiting it in another. While we found a strong treatment effect, i.e., explicit communication helps to coordinate actions such that strategic reports implement the preferred outcome of the majority, there is little evidence that tacit collusion works in the treatment without communication, despite the fact that all agents' preferences are common knowledge and there exists a simple symmetric collusive strategy for the majority. Individual dominant strategies are chosen by a proportion of one third (Gamma) to almost one half (Delta) of subjects in the majority, when extreme reports on the boundary of the strategy space could have ensured the majority's preferred outcome for Beta and Delta without tax payments. Looking for a behavioral approach that better explains why we do not observe tacit collusion of the majority in our data, we estimate a mixed quantal response equilibrium model (logit equilibrium). In this model, we assume that subjects either believe that errors are uncorrelated (QRE), or they believe that same-type subjects all use the same strategy (collusive QRE), and they best respond to a probability distribution of other players' strategies. This model shows that the minority has more incentives to submit extreme reports due to the specific loss functions in our experimental parametrization, hence managing to implement their preferred outcome which coincides with the social optimum. The estimated parameter for the proportion of subjects who believe that errors are uncorrelated decreases significantly when communication is introduced, hence showing that communication serves as a device to correlate strategies of same-type subjects.

Learning has no effect on the selection of a collusive outcome for the majority when explicit communication is not allowed. Only the introduction of communication has a strong effect on the outcomes: we observe over 60% of collusive outcomes in early rounds, and gaining experience further enhances collusion, so that in over 90% of late rounds the majority attains their payoff-maximizing outcome. M-type subjects learn to submit extreme reports which ensure this outcome at no risk of paying taxes only in the treatment with communication. Outcomes that would have implemented more equal total payoffs to all subjects have no importance here. While we chose the communication structure in this experiment such that it enhances collusion, it may be true that we also created an inter-group competition between majority and minority, which lead to such little concern for



the payoffs of the other subgroup. Future work could involve the question whether the information structure matters, e.g. whether the results would differ if we allow for communication among all group members, or if subjects can choose with whom to communicate.

## 6 References

- Attiyeh, G., Franciosi, R. and Isaac, M., 2000. Experiments with the pivot process for providing public goods. *Public Choice* 102, 95114.
- Aumann, R.J., 1959. Acceptable points in general cooperative n-person games. In: *Contributions to the Theory of Games IV* (A.W. Tucker and R.D. Luce, eds.), Princeton, N.Y.: Princeton University Press, 287-324.
- Bernheim, B.D., Peleg, B. and Whinston, M.D., 1987. Coalition-proof Nash equilibria. I. Concepts. *Journal of Economic Theory* 42, 1-12.
- Bornstein, G., Gneezy, U. and Nagel, R., 2002. The effect of intergroup competition on group coordination: an experimental study. *Games and Economic Behavior* 41(1), 1-25.
- Bosman, R., Hennig-Schmidt, H. and van Winden, F., 2006. Exploring Group Decision Making in a Power-to-Take Experiment. *Experimental Economics* 9, 3551.
- Burns, P., 1985. Market Structure and Market Behavior: Price Adjustment in a Multi-Object Progressive Oral Auction. *Journal of Economic Behavior and Organization* 6, 275-300.
- Burtraw, d., Goeree J., Holt, C. A., Myers, E., Palmer K. and Shobe, W., 2008. Collusion in Auctions for Emission Permits: An Experimental Analysis. *Journal of Policy Analysis and Management* 28(4), 672-691.
- Clarke, E.H., 1971. Multi-part Pricing of Public Goods. *Public Choice*, Vol. 11, 17-33.
- Charness, G. and Rabin, M., 2002. Understanding social preferences with simple tests. *Quarterly Journal of Economics* 117, 817-869.
- Charness, G., Rigotti, L. and Rustichini, A., 2007. Individual Behavior and Group Membership. *American Economic Review* 97(4), 1340-1352.
- Cason, T., Saijo, T., Sjöstrom, T. and Yamato, T., 2006. Secure implementation experiments: Do strategy-proof mechanisms really work? *Games and Economic Behavior* 57(2), 206-235.
- Dash, R. K., Parkes, D. C. and Jennings, N. R., 2003. Computational mechanism design: A call to arms. *IEEE Intelligent Systems*, 18(6), 4047.
- Dash, R. K., Rogers, A. and Jennings, N. R., 2004. A mechanism for multiple goods and interdependent valuations. In: *Agent-Mediated Electronic Commerce VI, Theories for and Engineering of Distributed Mechanisms and Systems, AAMAS 2004 Workshop, AMEC 2004, New York, NY, USA, July 19, 2004, Revised Selected Papers, Lecture Notes in Computer Science* pp. 15-29, Springer.
- Fehr, E. and Schmidt, K., 1999. A Theory of Fairness, Competition and Cooperation. *Quarterly Journal of Economics* 114, 817-68.
- Fischbacher, U., 2007. Z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics*, 10, 171-78.
- Goeree, J. K., Holt, C. A., and Palfrey, T. R., 2005. Regular Quantal Response Equilibrium. *Experimental Economics* 8, 347367.

- Green, J. R., and Laffont J.-J., 1977. Characterization of satisfactory mechanisms for the revelation of preferences for public goods, *Econometrica* 45(2), 427-438.
- Greiner, B., 2004. The Online Recruitment System ORSEE 2.0 - A Guide for the Organization of Experiments in Economics. University of Cologne, Working Paper Series in Economics 10.
- Groves, T., 1973. Incentives in Teams. *Econometrica* 41, 617-663.
- Groves, T. and Ledyard, J.O., 1977a. Limitations of Demand Revealing Processes. *Public Choice* 29, Special Supplement Devoted to the Demand-Revealing Process, 107-124.
- Groves, T. and Ledyard, J.O., 1977b. Reply to Comments by Tidemand and Tullock and Greenberg, Mackay and Tideman on Some Limitations of Demand Revealing Processes. *Public Choice* 29, Special Supplement Devoted to the Demand-Revealing Process, 139-143.
- Groves, T. and Loeb, M., 1975. Incentives and public inputs. *Journal of Public Economics* 4, 211-226.
- Haan, M. A., Schoonbeek, L. and Winkel, B. M., 2009. Experimental Results on Collusion. In: *Experiments and Competition Policy* (J. Hinloopen and H. Normann, eds.). Cambridge: Cambridge University Press, pp. 9-33.
- Isaac, R. M. and Walker, J. M., 1985. Information and Conspiracy in Sealed Bid Auctions. *Journal of Economic Behavior and Organization* 6, 139-59.
- Isaac, R. M. and Smith, V. L., 1985. In Search of Predatory Pricing. *Journal of Political Economy* 93, 320-345.
- Kawagoe, T. and Mori, T., 2001. Can the Pivotal mechanism induce truth-telling? An experimental study. *Public Choice* 108, 331-354.
- Kube, S. and Puppe, C., 2009. (When and how) Do voters try to manipulate? Experimental evidence from Borda elections. *Public Choice* 139, 395-2.
- Kwasnica, A. M. and Sherstyuk, K., 2007. Collusion and equilibrium selection in auctions. *Economic Journal* 117, 120-145.
- Li, J. and Plott, C., 2009. Tacit Collusion in Auctions and Conditions for its Facilitation and Prevention: Equilibrium Selection in Laboratory Experimental Markets. *Economic Inquiry* 47 (3), 425-448.
- McKelvey, R. D. and Palfrey, T. R., 1995. Quantal Response Equilibrium. *Games and Economic Behavior* 10, 6-38.
- Peleg, B., 1998. Almost all equilibria in dominant strategies are coalition proof. *Economics letters* 60, 157-162.
- Sherstyuk, K., 1999. Collusion Without Conspiracy: An Experimental Study of One-Sided Auctions. *Experimental Economics* 2, 59-75.
- Tideman, N. and Tullock, G., 1977. Some Limitations of Demand Revealing processes: Comment. *Public Choice* 29, Special Supplement Devoted to the Demand-Revealing Process, 125-128.
- Tideman, T.N., 1983. An experiment in the demand-revealing process. *Public Choice* 41, 387-401.

Vickrey, W., 1961. Counterspeculation, Auctions, and Competitive Sealed-Tenders. *Journal of Finance* 16, 8-37.

Weizsäcker, G., 2003. Ignoring the rationality of others: evidence from experimental normal-form games. *Games and Economic Behavior* 44, 145-171.

## **7 Appendix: Payoff Tables**

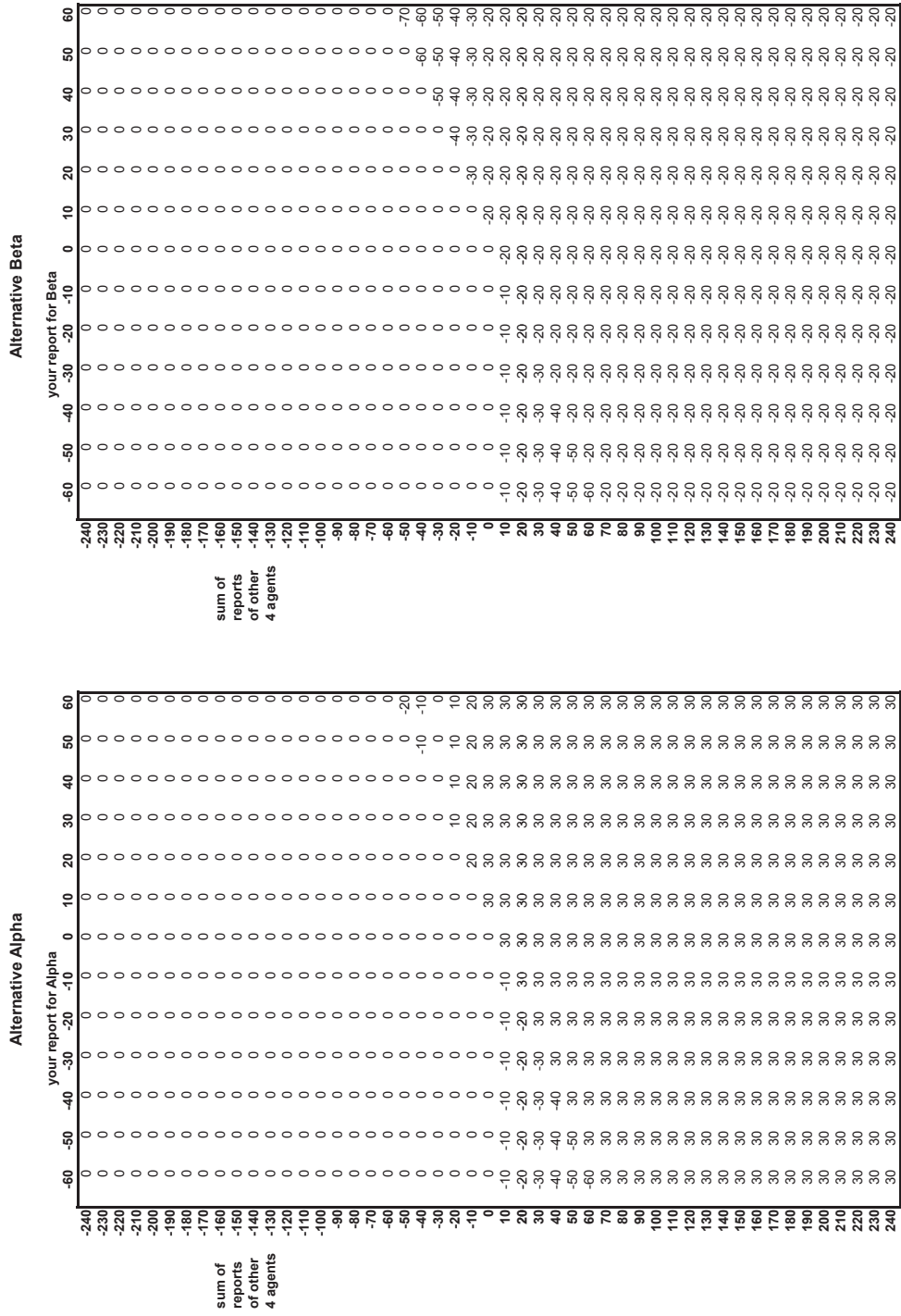


Figure 7: M's Payoff Tables for Alternatives Alpha and Beta

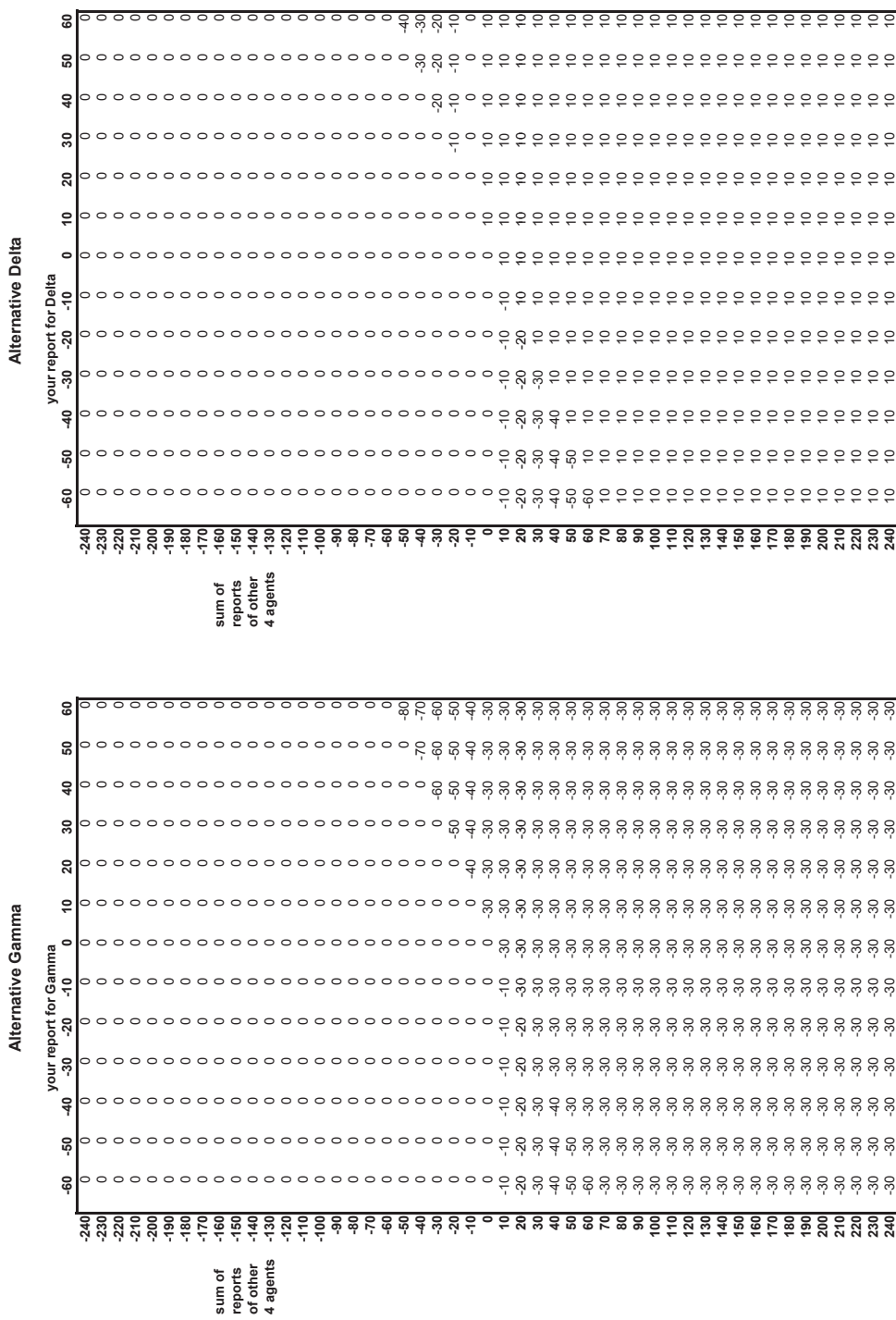


Figure 8: M's Payoff Tables for Alternatives Gamma and Delta

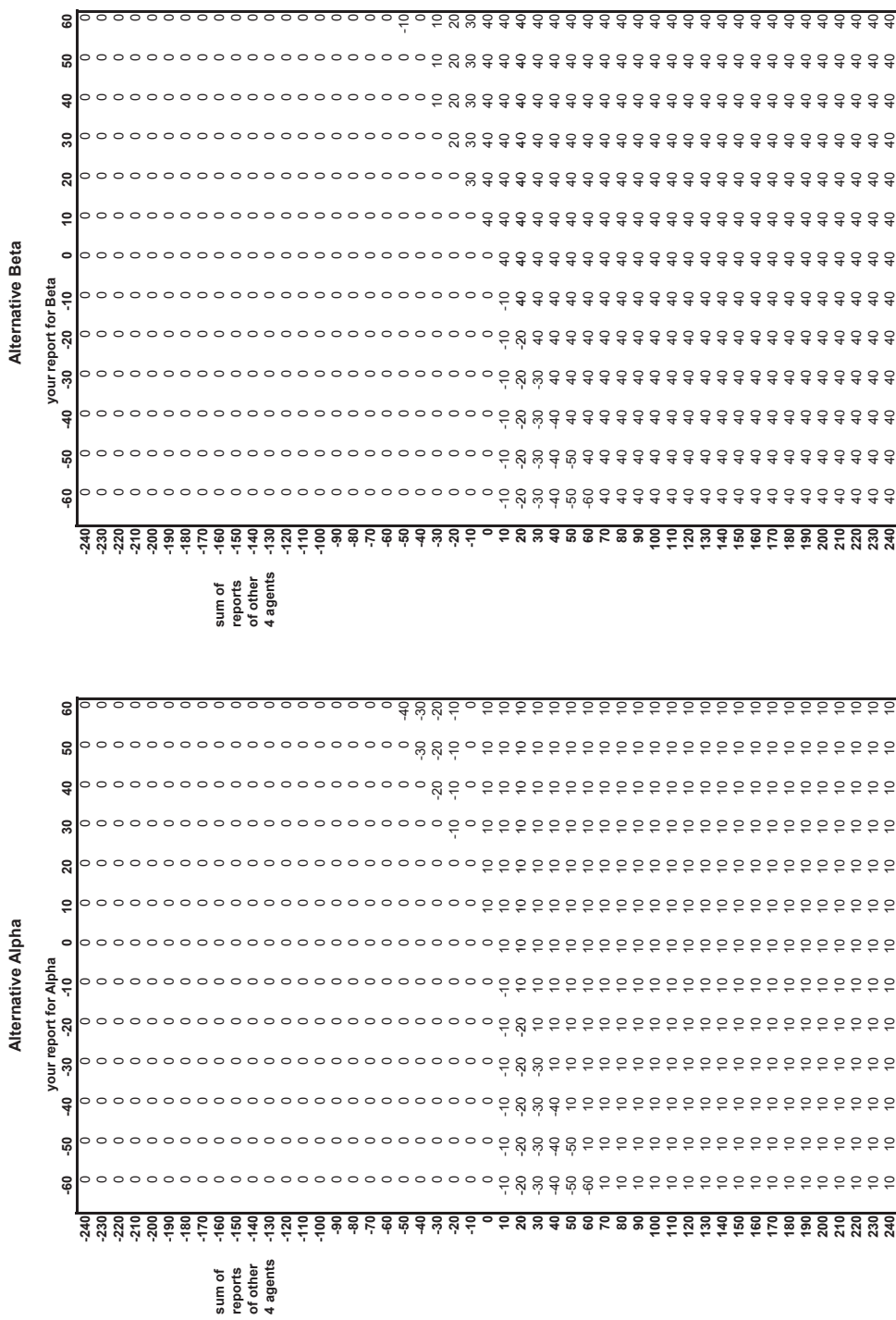


Figure 9: N's Payoff Tables for Alternatives Alpha and Beta

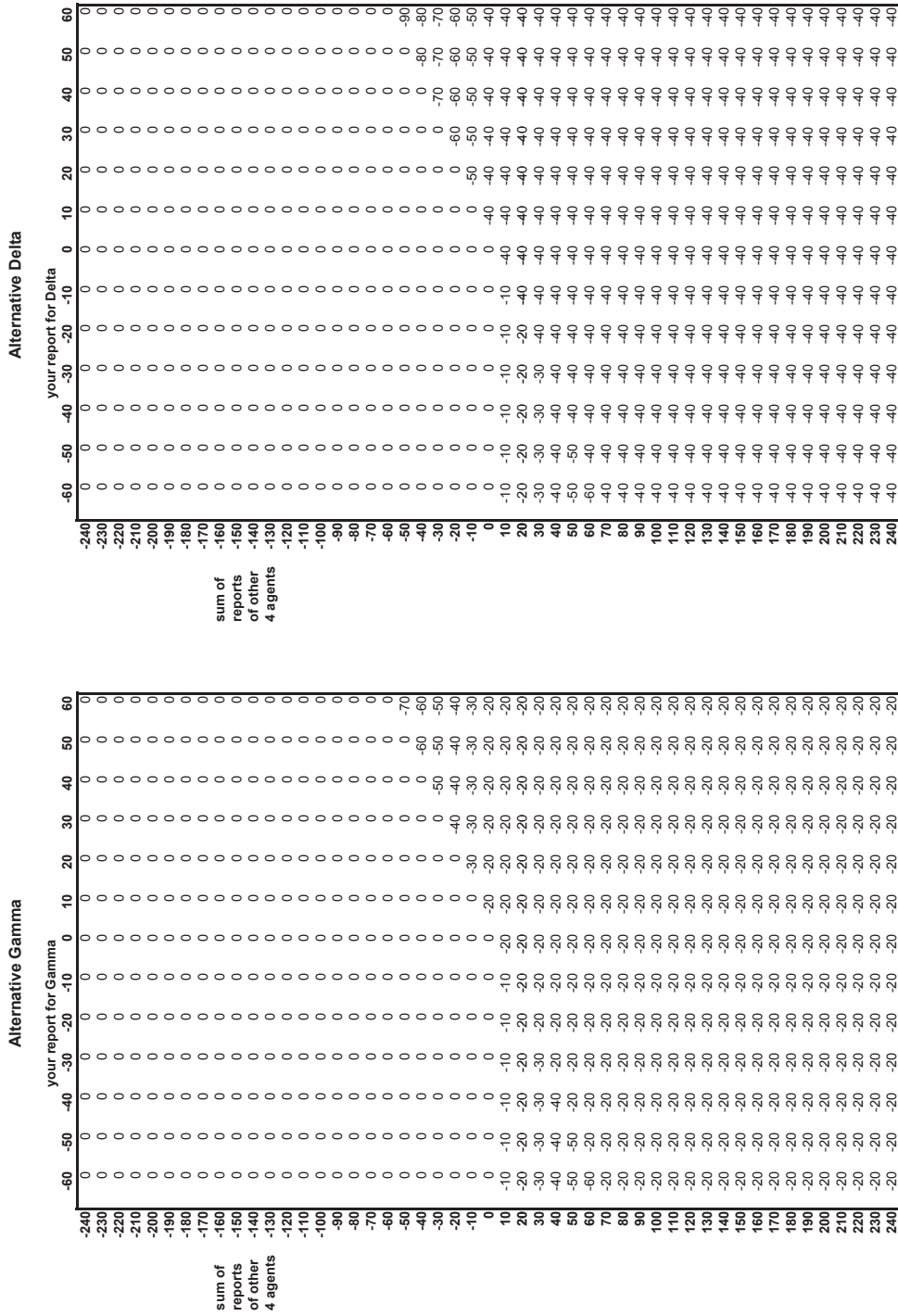


Figure 10: N's Payoff Tables for Alternatives Gamma and Delta