Online Appendix for: Responsibility Attribution for Collective Decision Makers*

Raymond Duch
Nuffield College Oxford

Wojtek Przepiorka
University of Oxford

Randolph Stevenson
Rice University †

April 27, 2014

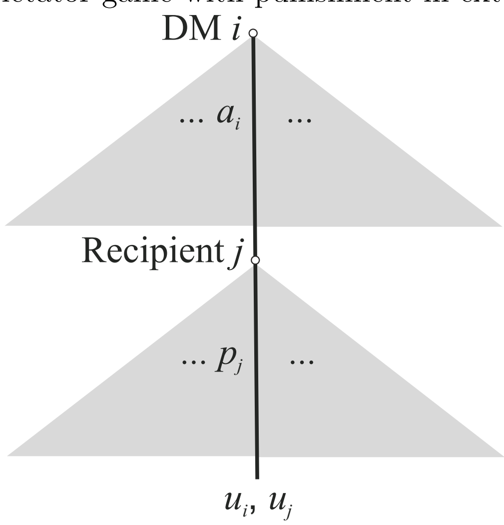
*The experiments reported in this paper were conducted at Nuffield CESS (Nuffield College Centre for Experimental Social Sciences). This work was partly supported by generous funding from Nuffield College Centre for Experimental Social Sciences and the Swiss National Science Foundation [grant number 100017_124877]. We thank Inaki Sagarzazu and Akitaka Matsuo for their excellent research assistance. We would like to thank Yuval Heller and Dominik Duell for their very helpful comments. Finally we are very much indebted for the comments received by the participants of the Comparative Politics Seminar at Stanford University, Politics Seminar at the University of Essex, the LSE Politics Seminar, the Economics Seminar Series at the University of Cape Town, South Africa, the University of Oxford Economics Society, the Cambridge Society for Economic Pluralism, the Hertie School Seminar Series and the Texas A&M Political Science Seminar. Replication data and files are available at http://www.raymondduch.com/replications.

†Raymond Duch is an Official Fellow at Nuffield College and Director of the Centre for Experimental Social Sciences, Nuffield College OX1 1NF Oxford (raymond.duch@nuffield.ox.ac.uk). Wojtek Przepiorka is a Post Doctoral Research Fellow in the Sociology Department and at the Nuffield Centre for Experimental Social Sciences, Nuffield College OX1 1NF Oxford (wojtek.przepiorka@sociology.ox.ac.uk). Randolph Stevenson is Professor of Political Science at Rice University P.O. Box 1892 Houston, TX 77251-1892 (randystevenson@rice.edu)

1 Game theoretic analysis

This section sketches a game theoretic analysis of the extensive-from game depicted in Figure 1. It is a simplified representation of the game used in our experiment as it is described as one between one DM i and one recipient j. In this game, the DM moves first and decides how much to keep for themselves and how much to allocate to the recipient. The DM can choose to keep a_i, where $a_i \in 0, 1, 2, \ldots, 25$ (out of $n = 25$). Consequently, the recipient is allocated $n - a_i$, that is 25, 24, 23, ..., or 0, respectively. The recipient observes the DM's move and can choose to punish the DM with p_j deduction points, where $p_j \in 0, 1, 2, \ldots, 600$. Since we assume exclusively self-regarding actors, the DM's utility u_i corresponds to their total payoff $\pi_i = \frac{a_i}{5} - \frac{0.1}{n-5} p_j$, and the recipient's utility u_j corresponds their payoff $\pi_j = \frac{n-a_i}{n-5}$ (see the main paper for the intuition behind this payoff function).

Figure 1: Dictator game with punishment in extensive form



This game is a complete-information sequential game and can thus be solved by backward induction. Starting from the bottom, we can first note that the recipient has no dominant action for any of the choices of the DM because the recipient's payoff π_j is not affected by their choice of p_j. In other words, whatever the DM does, the recipient is indifferent with regard to the amount of punishment to be inflicted on the DM. The DM, however, has a clear preference for choosing $a_i = 25$ because, irrespective of what the recipient does thereafter, this choice maximizes the DM's payoff. Since the recipient is indifferent as to whether and how much to punish the DM, the game has many Nash equilibria. However, we designed our experiment with the intuition that recipients would not be entirely indifferent between actions but rather exhibit some degree of other-regarding preferences; assuming that actors are slightly inequity-averse, eliminates the problem of indifference and multiple equilibria.

As we point out in the text of the article (and repeat here for the reader), one way to formalize inequity aversion is the Fehr-Schmidt model (**??**), in which a DM's utility depends on their payoff and the payoff of the recipient in the following way:

$$u_i = \begin{cases} \pi_i & \text{if } \pi_i = \pi_j \\ \pi_i - \alpha_i(\pi_j - \pi_i) & \text{if } \pi_i < \pi_j \\ \pi_i - \beta_i(\pi_i - \pi_j) & \text{if } \pi_i > \pi_j \end{cases} \tag{1}$$

DM i's payoff is denoted as before by π_i, α_i is the so called 'envy' parameter, and β_i the so called 'guilt' parameter. It is assumed that $\alpha_i \geq 0$, $\alpha_i \geq \beta_i \geq 0$ and $\beta_i < 1$. That is to say, no DM i likes having less than recipient j, no DM i likes having more than recipient j but dislikes having less more than he or she dislikes having more, and DM i's dislike for having more never outweighs his or her utility from having π_i, respectively. The recipient's utility function can be written accordingly by interchanging subscripts.[1]

We now continue our analysis whilst assuming that the DM and the recipient are Fehr-Schmidt rational with $\alpha_i = \alpha_j = 0.01$ and $\beta_i = \beta_j = 0.01$ (or any arbitrarily small values) and, as before, we start from the bottom by looking first at the recipient's actions conditional on the DM's actions. Recall first that a recipient's choice of $p_j > 0$

[1]There are other models of other-regarding preferences which could be used instead of the Fehr-Schmidt model, and which, with all likelihood, would yield qualitatively similar results (e.g. **???**). We use the Fehr-Schmidt model here because of its simplicity and intuitive appeal.

only reduces the DM's payoff and leaves the recipient's payoff unaffected. Since for any positive β_j the recipient prefers having the same payoff as the DM over the DM having a lower payoff, the recipient will never choose a p_j to obtain $\pi_i < \pi_j$ or to increase the payoff difference. More specifically, after replacing π_i and π_j with the right-hand side of the corresponding payoff functions and solving the equation for p_j, we can see that the recipient's choice of p_j will never exceed $50a_i - 250$.

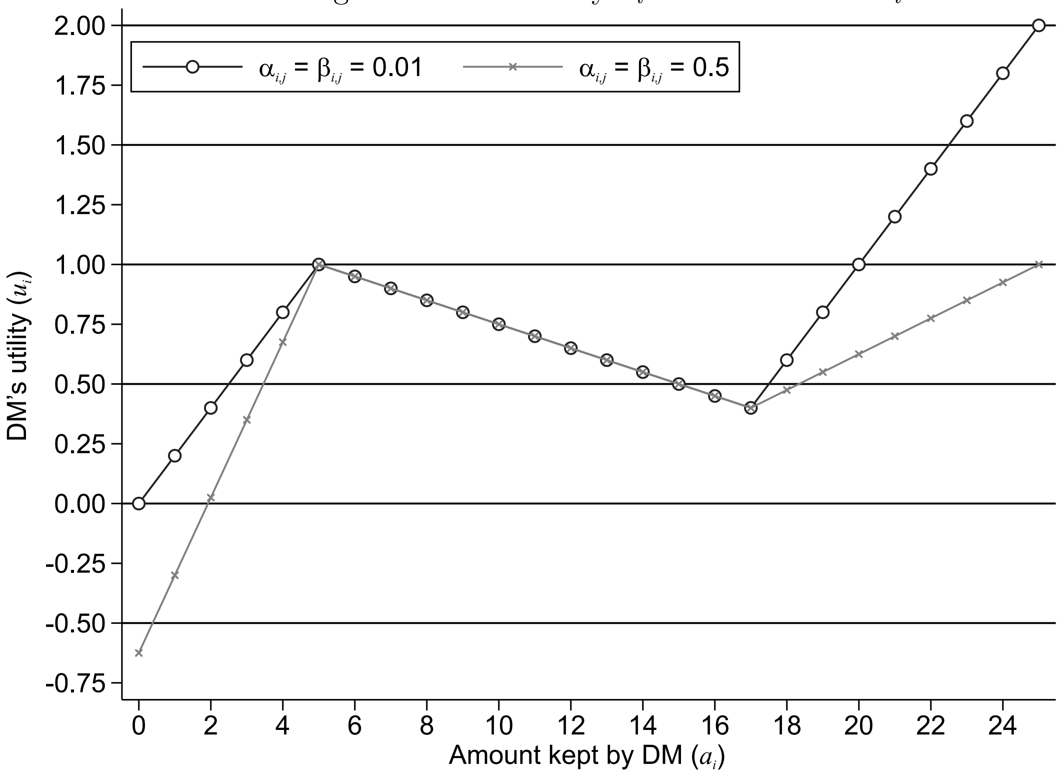
This has two important implications. First, the recipient will never punish an equitable or favourable allocation. That is, the recipient will never choose $p_j > 0$ if the DM chooses $a_i \leq 5$. Second, the recipient will inflict less than the maximum punishment on the DM (i.e. $p_j < 600$) if the DM chooses $a_i < 17$. Thus, for any allocation of 17 or more for the DM, the recipient will punish the DM maximally with $p_j = 600$. Now, as we know what the recipient will do conditional on the DM's actions, we can turn to the DM's actions.

First note that for any positive α_i the DM will never choose an allocation that favours the recipient. Second, any allocation $17 \geq a_i \geq 6$ will be punished by the recipient such that the DM's payoff will be reduced to the payoff of the recipient and thus an equitable outcome (i.e. $\pi_i = \pi_j$) will result. However, since the recipient's payoff decreases with the amount the DM keeps for themselves, so does the DM's payoff and reaches a local minimum at $a_i = 17$; the local maximum is at the equitable allocation $a_i = 5$. Finally, any allocation $a_i > 17$ will be punished by $p_j = 600$ and thus the maximum allocation $a_i = 25$ will result in the highest net payoff for the DM. Since the net payoff from choosing $a_i = 25$ is larger than the payoff from choosing $a_i = 5$, and the disutility the DM experiences from being minimally inequity avers is still negligible, the DM maximizes their utility by choosing $a_i = 25$. This identifies the unique subgame perfect Nash equilibrium in this game.[2] In this equilibrium, the DM chooses $a_i = 25$ and the recipient chooses

[2]Subgame perfection is an equilibrium refinement concept in game theory that excludes Nash equilibria in which players' responses to out-of-equilibrium behavior are not utility maximizing.

$p_j = 50a_i - 250$ subject to $600 \geq p_j \geq 0$. Figure 2 illustrates how the DM's utility changes as a function of a_i.

Figure 2: DM's utility u_i as a function of a_i



It is now easy to see that as the DM becomes increasingly inequity averse, the payoff maximizing action (i.e. $a_i = 25$) looses its appeal as the resulting outcome in terms of utility approaches the outcome of the equitable allocation (i.e. $a_i = 5$). Eventually, with $\beta_i = 0.5$, the DM becomes indifferent between keeping 25 and being maximally punished, and keeping only 5 and not being punished at all (see Figure 2). Now the game has two subgame perfect Nash equilibria. However, the DM is indifferent between the two actions only for $\beta_i = 0.5$. For $\beta_i > 0.5$ the unique subgame perfect equilibrium is the one in which the DM chooses the equitable allocation and the recipient does not punish. Note that this result is unaffected by an increase in α_i or α_j and β_j.

1.1 Discussion

Our ability to identify equilibrium choices in our experiment is greatly facilitated by assuming a single DM and a single recipient along with the assumption of very limited other-regarding preferences. Under these assumptions, there is a single equilibrium in which the DM allocates everything to themselves and the recipient expends all of their 600 punishment points to punish the single DM. Recall from our analysis above that in equilibrium, it is not possible that the recipient over-punishes the DM, and even in case of an out-of-equilibrium allocation by the DM, the recipient will calibrate the amount of punishment in order not to over punish the DM. In particular, for allocations $17 > a_i > 5$, the recipient will use $p_j = 50a_i - 250$ punishment points. In what follows, we will discuss some implications of relaxing these assumptions on recipients' punishment behavior.

Note first that the utility function as stated in **?**, 822 is:

$$u_i = \pi_i - \alpha_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_j - \pi_i; 0\} - \beta_i \frac{1}{n-1} \sum_{j \neq i} \max\{\pi_i - \pi_j; 0\}, \tag{2}$$

where n denotes the number of actors involved in the game (for $n = 2$, Equation 2 can be simplified to Equation 1). Let us now consider a recipient i's Fehr-Schmidt utility if $n = k + m$, where k denotes the number of DMs (e.g., $k = 5$) and m denotes the number of recipients (e.g., $m = 20$). Moreover, let's assume that DMs have agreed on an allocation that favours the DMs (i.e., $a > 5$). Now, a recipient i's utility can not only be affected by their payoff difference with DMs, but also by their payoff difference with other recipients as well as the group size n.

Note that in our collective dictator game, every recipient's payoff is always the same. Thus, recipient i's utility will in fact be unaffected by other recipients' payoffs. It will, however, be affected by DMs' payoffs. Since recipient i's payoff difference with DMs is the same for every DM, recipient i's disutility caused by these payoff differences will be $\alpha_i \frac{k}{k+m-1}(\pi_j - \pi_i)$, where j is a DM. Now we can see that as the number of recipients

m increases, a recipient i's disutility from having less than the DMs decreases. In other words, the more recipients are affected by an allocation decision made by the DMs, the smaller will be a single recipient i's disutility from this inequitable allocation.

Although the number of recipients m scales down a recipient i's 'envy' parameter α_i in the general Fehr-Schmidt model, it does not change our predictions regarding recipients' punishment behavior as these predictions hold for any arbitrarily small values of α_i. With regard to DMs behavior, the predictions based on the general model become more conservative. A DM i's disutility from allocating less to recipients is now $\beta_i \frac{m}{k+m-1}(\pi_i - \pi_j)$, where j is now a recipient. Thus, a DM i's 'guilt' parameter β_i is scaled down by the number of DMs k. As a consequence, the β-threshold at which the second subgame perfect Nash equilibrium emerges, where a DM is indifferent between choosing $a = 25$ and being maximally punished and choosing $a = 5$ and not being punished at all, will be higher, namely at $\frac{k+m-1}{2m}$ (which is 0.6 in our example). With this in mind, let us take a step back and first relax the assumption of a single DM and consider the implications of multiple DMs on the punishment behaviour of a single recipient.

In the case of k DMs (e.g., $k = 5$) and a single recipient, a Fehr-Schmidt rational recipient i maximizes the following utility function:

$$u_i = \pi_i - \alpha_i \frac{1}{k} \sum_{j \neq i} \max\{\pi_j - \pi_i; 0\} - \beta_i \frac{1}{k} \sum_{j \neq i} \max\{\pi_i - \pi_j; 0\}, \tag{3}$$

In this case again, if DMs keep the entire amount for themselves, the recipient cannot over-punish any DM, because by spreading their punishment points amongst multiple DMs (rather than concentrating them on a single DM), a DM's payoff cannot be reduced further (i.e., rather than receiving the maximum 600 point punishment possible in our experiment, in the spreading scenario, the DM will receive less than 600 punishment points). Moreover, since the inequity aversion parameters in the Fehr-Schmidt model are assumed to be linear, the recipient is indifferent between spreading their punishment points amongst multiple DMs and concentrating their punishment on a single DM (the sum of the reductions in inequity due to reducing each DMs payoff will equal the reduction from concentrating punishment on a single DM). Finally, since now punishment can be spread over multiple DMs, the recipient will be able to expend the maximum of 600 punishment points for a larger range of allocations by DMs. That is, for $k = 5$ DMs, the total amount of punishment must not exceed $5 \times (50a - 250)$. This implies that the single recipient will be able to expend the total of 600 punishment points for any allocation a, where $25 \geq a \geq 8$. For allocations of $a = 6$ or $a = 7$, the recipient will expend less than the 600 punishment points and no punishment at all for allocations of $5 \geq a \geq 0$.

We can further relax assumptions in our simplified game theoretic model so that rather than having a single recipient make all of the punishment decisions, the punishment decisions are taken by m (e.g., $m = 20$) recipients. The most serious implication of having multiple recipients is that it introduces the issue of coordination amongst recipients in allocating punishment points to individual DMs. Our collective dictator game with punishment is a one-shot game without communication and hence there is no opportunity for recipients to coordinate their punishment strategies. Hence, for DM allocations of $17 > a > 5$, the absence of coordination in our game theoretic model with other-regarding preferences raises the possibility of over or under punishment of certain DMs.

Finally, our simplified model does not accommodate individual DM decisions being conditioned on other-regarding assessments of other DMs' earnings. This of course is only relevant for DM payoffs that are adjusted by punishments (since they equally share the allocations they agree upon). One might imagine a version of our game theoretic model in which proposers (or voting DMs) anticipate punishment strategies and hence condition their proposals (or votes) on such anticipated punishments.

2 Experiment 1: Responsibility Attribution

2.1 Procedure and Design

We conducted one pilot session with 25 subjects and four experimental sessions with 25 subjects in sessions 1 and 4 and 23 subjects in sessions 2 and 3. Subjects were 29 years old on average (standard deviation = 14.1) and 45% were female. All subjects received £4 for showing up and earned £14.1 on average in the experiment. Upon arrival in the lab, subjects were given instructions on paper and the experimenter read the instructions aloud. Subjects then took a quiz comprised of eight control questions about the instructions. Correct answers appeared on the screen; questions with at least one wrong answer were read aloud by the experimenter and the correct answer was explained to all subjects. Next, the experiment was conducted. At the end of the experiment, subjects learned how much they had earned and were asked to fill in a questionnaire. Figures 3 through 6 reproduce the experimental instructions that subjects received in sessions 2, 3 and 4. The instructions for session 1 differed in several aspects (see below). All instructions contained screen shots of the proposing DM's money allocation screen (Figure 5), the DMs' voting screen, and the recipients' deduction screen (Figure 6). The reminder of this section describes the experimental design in more detail. Note that the information provided here is mostly complementary to what is provided in the main paper.

The pilot session differed from the experimental sessions in the following respects (Table 1). First, the four voting weight distributions differed. Second, it lasted for 24 rather than 20 rounds. Third, in each round, the DM with the largest vote weight was also the proposing DM. Fourth, we varied whether or not recipients knew what the vote weight distribution was in a round (this implies that they knew or did not know who the proposing DM was, respectively). Fifth, recipients were given 50 rather than 30 deduction points. Finally, for each deduction point a recipient did not use, he or she earned £0.01. In other words, punishment was costly in the pilot session. Since the session lasted longer than the 90 minutes we had scheduled and subjects were rather reluctant to spend their deduction points on punishing DMs, we reduced the number of rounds to 20, reduced the number of deduction points to 30 and made deductions costless in the following sessions. We also changed the voting weight distributions to achieve a more gradual variation.

The first session differed from sessions 2, 3 and 4 in that in session 1, we varied whether or not recipients knew what the vote weight distribution was (again, this implies that they knew or did not know who the proposing DM was, respectively). The results from the first session clearly indicated that recipients, if they did not know the vote weight distribution in a round, displayed two patterns of behavior: They either refrained from punishing DMs or punished each DM equally. Given this clear result, we decided to inform recipients about the vote weight distribution in each round in sessions 2, 3 and 4. Moreover, in session 1, the proposing DM was always the plurality DM. To disentangle the plurality-DM effect from the proposing-DM effect, we also varied the vote weight of the proposing DM in sessions 2, 3 and 4. Finally, we also varied whether or not recipients knew which one of the DMs the proposing DM was.

The results reported in sections 2.2.1 and 2.2.4 of the main paper are based on all data from sessions 1 through 4. The results reported in sections 2.2.2 and 2.2.3 of the main paper are based on the data from sessions 2, 3 and 4 where recipients knew which one of the DMs the proposing DM was (see Table 1).

Table 1: Experimental Design

Pilot session: 25 subjects, 24 rounds, largest DM is proposer, costly deduction

	Recipients know weights		Recipients know proposer	
Vote weights	yes	no	yes	no
51, 20, 15, 10, 4	4, 80	2, 40	4, 80	2, 40
49, 20, 15, 10, 6	4, 80	2, 40	4, 80	2, 40
35, 33, 32, 0, 0	4, 80	2, 40	4, 80	2, 40
22, 21, 20, 19, 18	4, 80	2, 40	4, 80	2, 40

Session 1: 25 subjects, 20 rounds, largest DM is proposer, costless deduction

	Recipients know weights		Recipients know proposer	
Vote weights	yes	no	yes	no
53, 29, 10, 6, 2	3, 60	2, 40	3, 60	2, 40
48, 19, 14, 11, 8	3, 60	2, 40	3, 60	2, 40
38, 21, 17, 13, 11	3, 60	2, 40	3, 60	2, 40
23, 21, 20, 19, 17	3, 60	2, 40	3, 60	2, 40

Sessions 2 - 4: 71 subjects, 3×20 rounds, random DM is proposer, costless deduction

	Recipients know weights		Recipients know proposer	
Vote weights	yes	no	yes	no
53, 29, 10, 6, 2	5, 280	0, 0	3, 168	2, 112
48, 19, 14, 11, 8	5, 280	0, 0	3, 168	2, 112
38, 21, 17, 13, 11	5, 280	0, 0	3, 168	2, 112
23, 21, 20, 19, 17	5, 280	0, 0	3, 168	2, 112

The table lists the number of rounds and the number of recipients' deduction decisions per treatment (one deduction decision corresponds to an allocation of deduction points by one recipient to the five DMs). The experimental design is a 4(vote weight distribution)×2(information condition) factorial within-subject design.

Figure 3: Instructions: Page 1

Dear participants,

Welcome and thank you for participating in this experiment. Before we describe the experiment, we wish to inform you of a number of rules and practical details.

Important rules

- The experiment will last for about **90 minutes**.
- Your participation is considered **voluntary** and you are free to leave the room at any point if you wish to do so. In that case, we will only pay you the show-up fee of £4.
- **Silence:** Please do remain quiet from now on until the end of the experiment. Those who do not respect the silence requirement will be asked to leave the experimental room. You will have the opportunity to ask questions in a few minutes.

What will happen at the end of the experiment?

Once the experiment is finished, please remain seated. We will need around 10 minutes to prepare your payment. You will be called up successively by the number on your table; you will then receive an envelope with your earnings and you will be asked to sign a receipt.

Description of the experiment

This experiment consists of 20 rounds**.**

Up to 25 subjects are participating in this experiment. Your earnings from the experiment depend on both the decisions you will make and the decisions other subjects will make. You will receive your payment after the completion of the 20 rounds of the experiment.

In each round of this experiment you, along with the other participants, will be randomly assigned to be either one of 5 “decision-maker” or a “recipient”. Decision-makers will be asked to decide how to allocate a fixed sum of money to each recipient.

After the decision-makers decide on the allocation, the recipients will review the allocation decisions. Recipients can decrease the decision makers’ pay-off by assigning deduction points to each decision maker or leave their pay-offs unchanged.

The next section describes the decision situation in more detail.

Figure 4: Instructions: Page 2

Course of one round of the experiment

1. Five decision-makers will be chosen at random. Depending on the number of participants, they will be given up to £25.

2. The five decision-makers will be assigned a specific set of votes – that is, 100 votes will be split at random among the five decision-makers (e.g., 42, 10, 28, 5, and 15).

3. (see Figure 1) One of the decision-makers who will be chosen at random will then propose an allocation of the money between the decision-makers and the recipients (for example, keep £10 for decision-makers and give £15 to recipients). The allocated amounts to the decision-makers and recipients are split equally among each group. (In the above example, each decision-maker would get £2 and if there were 20 recipients each would get £15/20 = £0.75).

4. (see Figure 2) Each of the five decision-makers then cast all their votes for the allocation or all against. Abstention is not allowed. If the allocation receives at least 51 votes then it passes. If it does not, then the proposer proposes a different allocation and it is also voted on. No other communication is allowed.

5. If the proposer makes 3 unsuccessful proposals (that is, fails to obtain 51 votes for any of his three proposals), the group will be disbanded and no one will be paid anything for that round. A new group of five decision-makers will be chosen randomly in the next round.

6. Information about the decision will be presented on recipients' screens. This will include the amount allocated to the decision-making group and to the recipients and the number of votes (out of 100) that each of the five decision-makers had to cast. In some rounds recipients will learn which of the five decision-makers proposed the allocation and in some rounds this information will not be shown.

7. Following the decision by the decision-makers, the recipients will review the allocation and decide whether (and how much) to decrease the income of each decision-maker.

8. (see Figure 3) In each round, each recipient will be given 30 "deduction points" that he or she may keep or assign to individual members of the decision-making group. Recipients may assign all 30 points to one member, six to each, no deduction points to any member, or anything in between. Each recipient decides independently about the assignment.

9. A decision-maker's deduction for the round is the average of the deduction points assigned to him or her by all recipients times £0.1. (For example, if 10 recipients assign 15 deduction points to a decision-maker, five assign ten points and five zero points, then the assigned deduction is £0.1*(10*15+5*10+5*0)/20 = £1.)

10. A decision-maker's deduction will be subtracted from his or her pay-offs from the round. However, after each round the decision-makers will get no feedback on the amount of their deduction – they will only learn their total pay-off at the end of the experiment.

This process is repeated in each of the 20 rounds.

Figure 5: Instructions: Page 3

Page | **3**

Your total earnings

Your total earnings in this experiment comprise the amount you earned in allocations in each round minus £0.1 times the number of deduction points you received in each round. If the amount you earned in all rounds is less than zero, you will still get the £4 of the show-up fee (so the minimum you can earn is £4).

Please answer the control questions that appear on your screen. You may use these instructions to answer the questions.

Please leave these instructions on your table when you leave the room.

Figure 1: Decision-maker proposing an allocation

CONSIDER ATTENTIVELY

In this period, you are a

decision-maker

The voting weights in this period are

53, 29, 10, 6, 2

All participants know that the decision-maker with **29 votes** proposes the allocation.

You have been assigned a voting weight of

29

There are participants in this experiment. Please, make an allocation of £ . If your allocation receives at lest 51 votes, the allocated amounts will be split equally among the members of each group.

decision-makers get:

recipients get:

Figure 6: Instructions: Page 4

Figure 2: Decision-makers vote on allocation

The decision-maker with **29** votes has proposed the following allocation:

decision-makers get:	£	(£ per decision-maker)
recipients get:	£	(£ per recipient)

You can now decide whether or not to approve this allocation. Your decision will be given a weight of **10 votes**. The distribution of votes in this period is **53, 29, 10, 6, 2**.

If the allocation receives at least 51 votes, it will come into effect. Otherwise, the decision-maker with **29** votes can propose a new allocation.

APPROVE | NOT APPROVE

Figure 3: Recipients decide whether to assign deduction points

The decision-maker with **29 votes** has proposed the following allocation:

decision-makers get:	£	(£ per decision-maker)
recipients get:	£	(£ per recipient)

The proposed allocation **has been approved** with a majority of votes.

You have 30 deduction points. You can now decide whether (and how many) of your points you want to assign to each decision-maker. For each 10 points assigned to a decision-maker on average, his or her income will be reduced by £1.

Decision-maker:	1	2	3	4	5
Number of votes:	6	29	53	10	2
Your deduction:	0	0	0	0	0

2.2 Results: Multivariate Model

2.2.1 Further particulars of the estimation procedure

Our dependent variable is a five-part composition that gives the share of the total deduction points used on each DM by a recipient. We follow Atchison's (**?**) recommendations in estimating this as compositional data. First, we choose a baseline category (i.e., allocations to a specific DM) and then divide each of the other four allocations by this baseline. We then take the log of each ratio. Our dependent variable, in a seemingly unrelated regression (SUR), is the vector of the log of these four ratios. Once we obtain estimates of the parameters of this model, we then get predicted shares of punishment for all five DMs (for a given subject-period) by transforming back to the original shares via the multivariate logistic transformation (**?**). Finally, we add confidence intervals around our predictions and substantive effects using the simulation methods described in **?**. The correlation across DMs for a given recipient is accounted for in the estimation of the SUR (which estimates the correlation across equations). To address the possible correlation across periods for the same subject, we also estimated models with robust standard errors, clustered on subject.[3]

Since we cannot characterise shares of punishment over DMs when recipients punished no one, in the compositional analysis reported below, we use only those subject-periods in which some allocation of deduction points was made (35% of recipients allocated none of their deduction points while 55% allocated all of them). Thus, the results from these models should be interpreted as conditional on recipients in a period making some allocation. Of course, one can still model the initial decision to punish or not; but, as we show in the next section, this analysis reveals one overwhelmingly significant predic-

[3] In order to produce these estimates we first replicated the jointly estimated SUR models by estimating each equation separately. In this application, these coefficients are identical, though there are very small changes in the standard errors. Next, we clustered on subject in the separately estimated equations. This produced larger, but still very small differences in standard errors. These results are provided in Section 2.2.4 below.

tor of non-punishment: the generosity of the collective allocation from the DMs to the recipients.[4]

We estimate a seemingly unrelated regression that includes four equations (one for each log-ratio with the DM with the largest vote weight, DM 5, as the baseline category). Each equation includes indicators for the vote weight variables associated with each distribution. Also included are indicators of the proposer status of all the DM's. We also include a control in each equation for the total number of deduction points allocated in the period by the subject. This allows for the possibility that subjects change the distribution of their punishments as they punish more or less in total. Finally, a variable is included in the model that controls for the allocation decision by the DMs (the split between DMs and recipients). Since interpreting coefficients from compositional models with several sets of indicator variables can be difficult, we only present the substantive implications of the model in the main paper. The full set of estimates on which these substantive effects are based is provided in Section 2.2.4 below.

2.2.2 Recipients' decision to punish or not

As indicated in the previous section, the vast majority of our subjects used all or none of their deduction points. The whole distribution is given in Figure 7. In the main paper, we are focused on how those who used at least 1 deduction point distributed their points across DMs with different characteristics. Here we ask the prior question of what determined the choice to use no deductions points or some points (which usually meant using all one's points).

In Table 2, we give results from a simple specification with only one predictor that nonetheless predicts almost 80% of subject's decisions to punish or not. All other pre-

[4]Most recipients who punish do not punish all DMs. Consequently, some of the shares in each recipient-period's five-part composition are usually zero. Following **?**, we set zero shares (given they are not all zero) to a very small number (.00001) and then rescale to assure the composition still sums to one exactly).

dictors that we examined were insignificant and made no improvement in the fit of the model or the predictive accuracy. These included age, gender, indicators for the four distributions used in the treatments, income, and education level. Thus, we conclude that the decision to punish or not was overwhelmingly determined by the generosity or selfishness of the collective DMs.

Figure 7: Total punishment points allocated by recipients

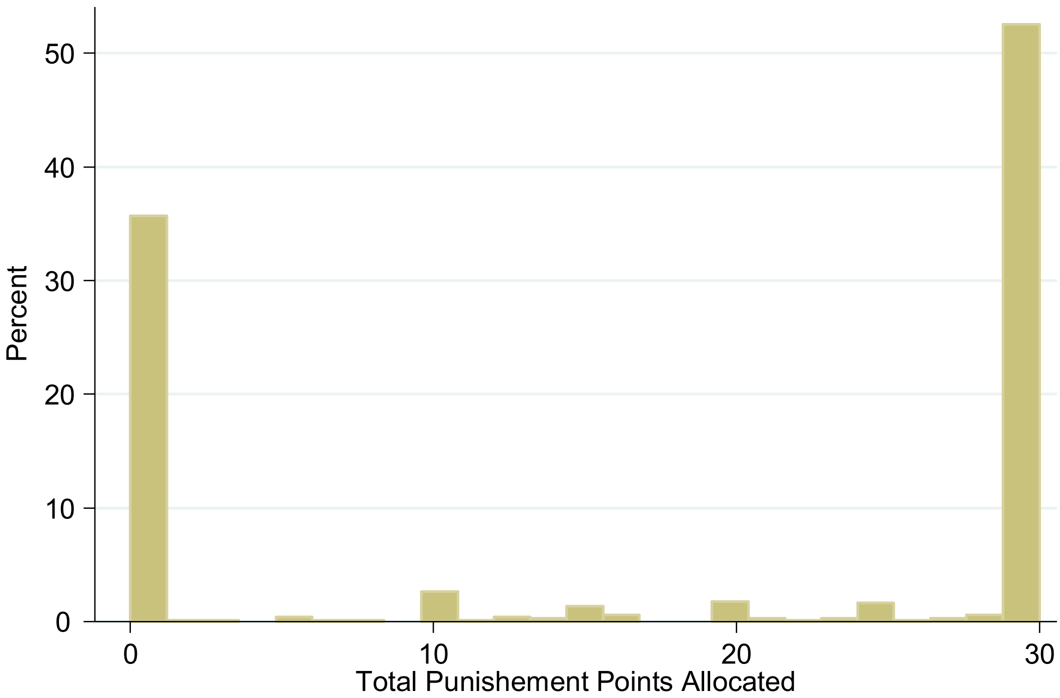


Table 2: Logit Regression of Decision to Punish

Dependent Variable: Deduction Points > 0	Coefficients (SE)	Effect of Change from £5 to £23 kept by DMs
Amount DMs kept	.26	62%
	(.04)	95% CI: (52.7%, 71.6%)
Constant	-.56	
	(.19)	
N_1 (observations)	912	
N_2 (clusters)	96	
Correctly predicted	77.96%	
Area under ROC curve	.803	

2.2.3 DMs' allocation decisions

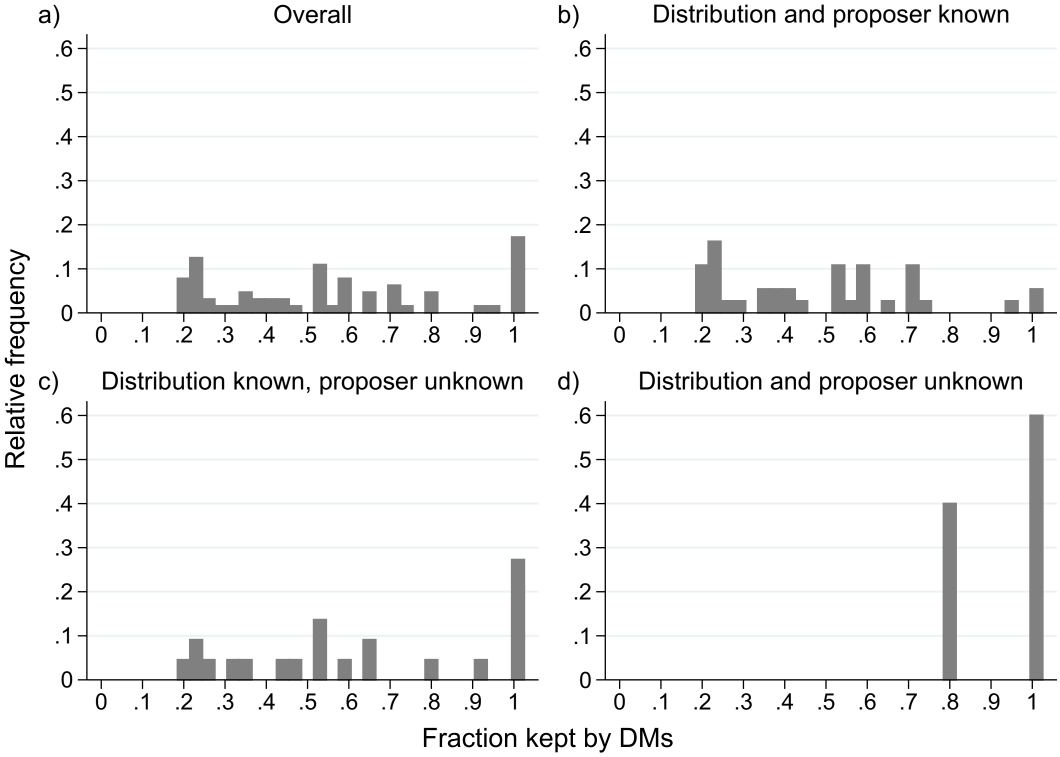
DMs' allocations are reasonably consistent with results from experiments with similar games. Our collective dictator game with second party punishment resembles the ultimatum game (?) because DMs may be concerned about offending recipients who can retaliate. In ultimatum games, in which there is a single proposer and a single recipient, offers typically amount to about 40 percent of the total endowment (?).

In spite of the punishment threat, DMs are not particularly generous in our game. Overall, only 20 percent of the allocations between DMs and recipients are equitable (Figure 8a).[5] Given that the allocations are divided amongst 5 DMs and 20 recipients, it makes also sense to compare the average division of allocations between a single DM and a single recipient. On average the individual DM receives 80 percent of the allocation (i.e., the 5 DMs collectively receive roughly half the allocation which is £12.5 or £2.5 per DM) compared to the 20 percent allocated to the individual recipient (£12.5 divided up amongst the 20 recipients results in £0.625 per recipient). Hence, a comparatively large proportion of DMs in our experiment are opting for allocations that are "unfair" to recipients. Nevertheless, relatively few opt for the strict equilibrium choice in which DMs keep all of the allocation.

The data from Experiment 1 in our manuscript clearly suggest that DMs condition their offers on the anticipation of recipient punishment. We are able to compare DM behavior under three different information conditions: when recipients are told DM vote weights and proposer status (Figure 8b), told only the vote weights (Figure 8c), or told neither vote weights nor proposer status (Figure 8d). It turns out that when recipients know only the vote weights, DMs' overall allocations to recipients fall to 40 percent (on average), and when recipients know nothing that could distinguish DMs form one another,

[5]Here, an equitable allocation is one where DMs keep 5 for themselves and allocate the rest to recipients. Since the number of recipients varied across sessions (see Section 2.1 above), the equitable fractions in Figure 8 are thus either .2 or .22.

Figure 8: Fraction kept by DMs across information treatments



DMs' allocations to recipients fall to only 5 percent (on average). These results are consistent with what **?** find. In their study, they explicitly explore how the anticipation of recipient punishment strategies shapes DM offers in a collective dictator game, suggesting that DMs are much more selfish when they think recipients have no way to separate them from the other DMs.

2.2.4 Multivariate Estimations for Figures 3 and 4 in main paper

This section provides the coefficient estimates that produced the substantive results in Figures 3 and 4 in the main paper as well as some additional estimations to check robustness of the results.

In Table 3, DM 1 always stands for the smallest DM in terms of voting weight within a distribution of weights; DM 2 stands for the next largest DM, etc.; and DM 5, the baseline category, is always the largest DM. In the top panel of the table, we give the estimates for the full set of indicators on weights in each of the four equations. This we do to facilitate comparing the coefficients for indicators on different vote weights in the same distribution. Reading from left to right for a given distribution, if the coefficients become increasingly positive, this would support the notion that the vote weight correlated positively with the punishment of DMs. In Table 3, we shade those coefficients that are significantly different than the coefficient to their left. Consistent with the substantive effects reported in the main paper, there is only weak evidence of a vote weight effect. In three cases the coefficients are significantly more positive, and in a forth case the coefficient is actually significantly more negative.

We estimate a SUR that includes four equations (one for each log-ratio with the DM 5 as the baseline category). Each equation includes indicators for the vote weight variables associated with each distribution (Distributions 1 through 4) which exclude the most equal vote distribution (Distribution 4). The weights corresponding to each of the distributions and to each of the equations is indicated below the coefficients in brackets. Also included is an indicator of the proposer status of all the DMs (proposer is DM 1 through 4 – Proposer 5 is excluded).[6] In each equation, we also include a control for the total number of deduction points allocated in the period by a recipient (Total Deduction Points). This allows for the possibility that recipients change the distribution of their punishments as they punish more or less in total. Finally, a variable (Amount Allocated to DMs) is included in the model that controls for the allocation decision by the DMs (the split between DMs and recipients).

The results presented in Table 6 are re-estimates of Table 3, but using an equation by equation estimation instead of joint estimation. This approach also allows estimating robust standard errors, clustering on subjects. Clearly, the differences in standard errors across specifications are quite small.

[6]This allows, for example, a change in the vote weight of DM 2 to impact the vote weight allocation of DM 4 not only through the deterministic connection between allocations for DM 2 and DM 4 but also through the coefficient on the included variable.

The SUR regressions in Table 3 include as an explanatory variable the DM allocations. Including a post-treatment response as an explanatory variable might reopen the backdoor path that randomization among voting rules was designed to close off. Sensitivity analyses suggest that this likelhood does not affect the causal effect of the treatment effects described in the main text. We first note that the design of Experiment 2 allows us to confirm our initial causal claims. Because of the design of this experiment, the post-treatment variable, by construction, is effectively removed from the model. In our Experiment 2 we use the strategy method and subjects respond to a fixed set of alternative DM allocations – as a result, we do not include on the right hand side of the model post-treatment variables. And, reassuringly, the results for Experiment 2 are consistent with the findings from Experiment 1.

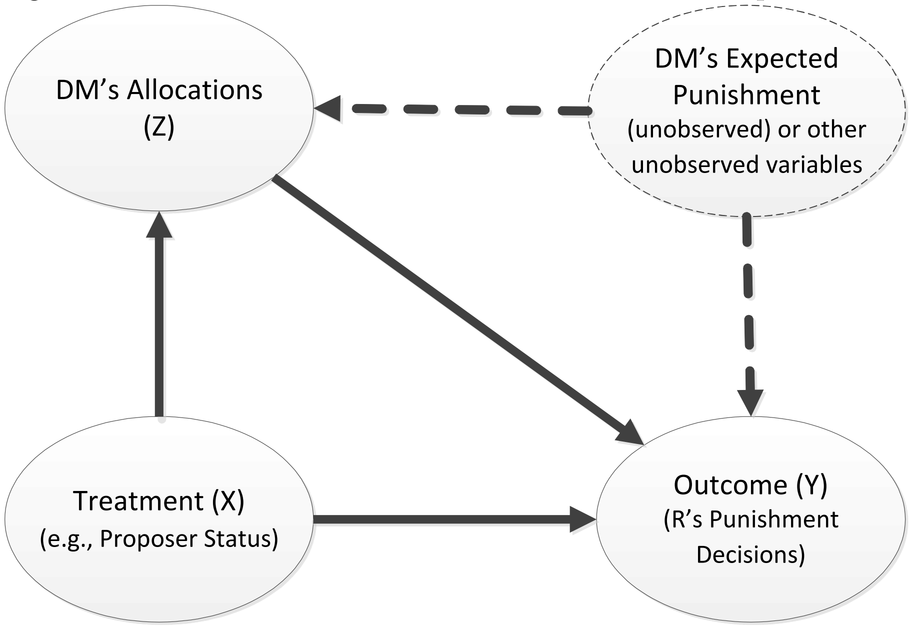
2.2.5 Causal structure of the model

It is also instructive to analyse, in more depth than the main text, the causal structure of the model that was estimated using the Experiment 1 data. Figure 9 summarizes the causal structure of the Experiment 1 model. This diagram indicates that our treatments (proposer status and voting weights) may have both a direct effect on recipient punishments as well as an effect mediated through DM allocations. Further, we should expect that DM allocations and recipient punishments share common (and unmeasured) causes (for example, expected punishments may well influence both actual punishments and DM allocations).

The situation pictured in Figure 9 corresponds exactly to example (b) in Figure 6 (page 681) of **?** that lists identified causal models. However, given the mediating status of the DM allocation variable, identification of the causal effects in this model may be more complicated than made clear in the original manuscript.

If Figure 9 correctly captures the causal relationships then the following holds:

Figure 9: Causal structure of treatment effects from Experiment 1



- The minimally sufficient set of controls for identifying the total effect of our treatments on punishments (i.e., both direct and mediating) does not include DM allocations
- The direct effect, of our treatment on punishment, by itself is not identified
- Including DM allocations re-establishes a backdoor path between our treatments and punishments.

The most obvious thing to do in this situation is to re-run all the models without the DM allocation variable. This means that the resulting estimate is the total causal effect of treatments on recipient punishments.

Accordingly we re-estimated the models in Table 3 without the DM allocation variable. The results of these re-estimated models are presented in Table 5. Some of the coefficients on the indicator variables for vote weights change but they do not in any fashion affect the substantive conclusions based on the original fully-specified model. And the coefficients on the proposer dummy variable are positive and statistically significant, again consistent with the results in the original model specifications.

Given the causal structure in Figure 9, this can only happen if one of the key connections there is not operative. We dismiss the possibility that DM allocations are not related to punishments (see Figure 1 in the main text). Likewise, while it could be that DM allocations are not connected (via common causes) to punishments, this seems unlikely given the large literature on ultimatum games (and games that share similar structures) that show DMs anticipate punishments and adjust their allocations accordingly.

Thus, the only remaining possibility is that our treatments do not impact DM allocations (that is we get the same DM decisions when we allocate different voting weights and procedural powers), so that there is no mediating effect. In this case, Figure 9 reduces to one with just X and Y and the direct effect equals the total effect, and one should get the same estimates of the effect whether one controls for DM allocations or not, as we do.

Table 3: Subject Punishments and Decision Maker Characteristics

	DM1/DM5	DM2/DM5	DM3/DM5	DM4/DM5
	Indicator Variables for Vote Weights (shaded cells are statistically different from cell to the left)			
Distribution 1	-3.64	-3.35	-4.58	-2.35
[Weight=53 excluded]	(-0.75)	(0.72)	(0.69)	(0.67)
	[Weight=2]	[Weight=6]	[Weight=10)	[Weight=29]
Distribution 2	-2.06	-2.57	-2.8	-2
[Weight=48 excluded]	(0.79)	(0.77)	(0.73)	(0.71)
	[Weight=8]	[Weight=11]	[Weight=14]	[Weight=19]
Distribution 3	-2.38	-1.33	-1.84	-0.26
[Weight=38 excluded]	(0.81)	(0.78)	(0.74)	(0.72)
	[Weight=11]	[Weight=13]	[Weight=17]	[Weight=21]
Distribution 4 (baseline)	–	–	–	–
[Weight=23]	[Weight=17]	[Weight=19]	[Weight=20]	[Weight=21]
	Indicator Variables for Proposer			
DM 1 is Proposer?	6.22	1.29	1.66	2.49
(1=yes 0=no)	(0.75)	(0.72)	(0.69)	(0.67)
DM 2 is Proposer?	2.13	5.61	2.19	2.23
(1=yes 0=no)	(1.15)	(1.11)	(1.05)	(1.03)
DM 3 is Proposer?	2.85	2.94	6.25	3.35
(1=yes 0=no)	(0.85)	(0.82)	(0.78)	(0.76)
DM 4 is Proposer?	1.71	2.44	2.49	5.11
(1=yes 0=no)	(0.89)	(0.86)	(0.82)	(0.80)
	Other Variables			
Amount Allocated to DMs	-0.15	-0.2	-0.24	-0.15
(0-25 points)	(0.07)	(0.07)	(0.07)	(0.07)
Total Deduction Points	0.1	0.11	0.03	0.03
(0-30 points)	(0.05)	(0.05)	(0.04)	(0.04)
Constant	-4.75	-4.57	-1.31	-2.95
	(1.62)	(1.60)	(1.52)	(1.48)
Obs.	433	433	433	433
RMSE	5.64	5.55	5.18	5.05
R^2	.22	.16	.27	.16

Table 4: Subject Deductions and DM Characteristics (further model estimations)

	M1		M2		M3	
	coef	se	coef	se	coef	se
DM1/DM5 [“Weight 17” and “DM 5 is proposer” are the reference categories]						
Weight 2	-3.639	0.747	-3.639	0.755	-3.639	0.760
Weight 8	-2.057	0.794	-2.057	0.803	-2.057	0.804
Weight 11	-2.378	0.806	-2.378	0.816	-2.378	0.813
DM 1 is proposer	6.221	0.749	6.221	0.758	6.221	0.829
DM 2 is proposer	2.135	1.146	2.135	1.160	2.135	1.115
DM 3 is proposer	2.852	0.848	2.852	0.858	2.852	0.887
DM 4 is proposer	1.710	0.891	1.710	0.902	1.710	0.926
Amount allocated to DMs	-0.146	0.073	-0.146	0.074	-0.146	0.082
Total deduction points	0.099	0.046	0.099	0.047	0.099	0.045
const.	-4.750	1.652	-4.750	1.671	-4.750	1.732
DM2/DM5 [“Weight 19” and “DM 5 is proposer” are the reference categories]						
Weight 6	-3.352	0.722	-3.352	0.731	-3.352	0.731
Weight 11	-2.566	0.768	-2.566	0.777	-2.566	0.750
Weight 13	-1.334	0.780	-1.334	0.789	-1.334	0.785
DM 1 is proposer	1.289	0.724	1.289	0.733	1.289	0.735
DM 2 is proposer	5.606	1.109	5.606	1.122	5.606	1.458
DM 3 is proposer	2.942	0.821	2.942	0.830	2.942	0.854
DM 4 is proposer	2.438	0.862	2.438	0.873	2.438	0.925
Amount allocated to DMs	-0.197	0.071	-0.197	0.072	-0.197	0.078
Total deduction points	0.106	0.045	0.106	0.045	0.106	0.044
const.	-4.574	1.598	-4.574	1.617	-4.574	1.705
DM3/DM5 [“Weight 20 and” “DM 5 is proposer” are the reference categories]						
Weight 10	-4.580	0.686	-4.580	0.694	-4.580	0.690
Weight 14	-2.797	0.730	-2.797	0.738	-2.797	0.697
Weight 17	-1.839	0.741	-1.839	0.750	-1.839	0.742
DM 1 is proposer	1.662	0.688	1.662	0.696	1.662	0.731
DM 2 is proposer	2.191	1.054	2.191	1.066	2.191	1.034
DM 3 is proposer	6.252	0.780	6.252	0.789	6.252	0.867
DM 4 is proposer	2.493	0.819	2.493	0.829	2.493	0.863
Amount allocated to DMs	-0.236	0.067	-0.236	0.068	-0.236	0.076
Total deduction points	0.034	0.043	0.034	0.043	0.034	0.057
const.	-1.312	1.518	-1.312	1.536	-1.312	1.973
DM4/DM5 [“Weight 21” and “DM 5 is proposer” are the reference categories]						
Weight 29	-2.348	0.668	-2.348	0.676	-2.348	0.686
Weight 19	-2.003	0.711	-2.003	0.719	-2.003	0.705
Weight 21	-0.262	0.722	-0.262	0.730	-0.262	0.676
DM 1 is proposer	2.493	0.670	2.493	0.678	2.493	0.736
DM 2 is proposer	2.231	1.026	2.231	1.038	2.231	1.033
DM 3 is proposer	3.347	0.759	3.347	0.768	3.347	0.775
DM 4 is proposer	5.110	0.798	5.110	0.807	5.110	0.887
Amount allocated to DMs	-0.150	0.066	-0.150	0.066	-0.150	0.075
Total deduction points	0.025	0.041	0.025	0.042	0.025	0.047
const.	-2.950	1.478	-2.950	1.496	-2.950	1.705

Notes: Model M1 is the same as the model in Table 3. Model M2 uses an equation by equation estimation instead of a joint estimation, and model M3 is the same as model M2 but lists robust standard errors, clustering on subjects.

Table 5: Subject Punishments and Decision Maker Characteristics (No DM Allocations in Model)

	DM1/DM5	DM2/DM5	DM3/DM5	DM4/DM5
	Indicator Variables for Vote Weights (shaded cells are statistically different from cell to the left)			
Distribution 1	-3.42	-3.06	-4.23	-2.13
[Weight=53 excluded]	(0.74)	(0.72)	(0.69)	(0.67)
	[Weight=2]	[Weight=6]	[Weight=10)	[Weight=29]
Distribution 2	-1.63	-1.99	-2.10	-1.56
[Weight=48 excluded]	(0.76)	(0.75)	(0.71)	(0.69)
	[Weight=8]	[Weight=11]	[Weight=14]	[Weight=19]
Distribution 3	-2.09	-0.96	-1.39	0.02
[Weight=38 excluded]	(0.80)	(0.78)	(0.74)	(0.72)
	[Weight=11]	[Weight=13]	[Weight=17]	[Weight=21]
Distribution 4 (baseline)	–	–	–	–
[Weight=23]	[Weight=17]	[Weight=19]	[Weight=20]	[Weight=21]
	Indicator Variables for Proposer			
DM 1 is Proposer?	6.48	1.63	2.08	2.75
(1=yes 0=no)	(0.74)	(0.72)	(0.69)	(0.66)
DM 2 is Proposer?	2.75	6.44	3.20	2.87
(1=yes 0=no)	(1.10)	(1.07)	(1.02)	(0.99)
DM 3 is Proposer?	3.44	3.73	7.20	3.95
(1=yes 0=no)	(0.80)	(0.78)	(0.74)	(0.72)
DM 4 is Proposer?	2.28	3.21	3.42	5.70
(1=yes 0=no)	(0.85)	(0.82)	(0.79)	(0.76)
	Other Variables			
Amount Allocated to DMs	-	-	-	-
(0-25 points)	-	-	-	-
Total Deduction Points	0.8	0.08	0.002	0.005
(0-30 points)	(0.05)	(0.04)	(0.04)	(0.04)
Constant	-6.38	-6.77	-3.95	-4.62
	(1.44)	(1.40)	(1.34)	(1.29)
Obs.	433	433	433	433
RMSE	5.66	5.50	5.25	5.08
R^2	.22	.14	.25	.15

2.2.6 More detailed (less aggregated substantive effects)

Figure 10 to Figure 13 give the separate substantive effects that were aggregated in Figure 3 in the main paper. Specifically, instead of averaging the share of punishment for all cases in which a DM of a given size was not the proposer (i.e., over cases when different other DMs were proposers), these figures give the estimates for each case separately. In addition, these graphs plot the raw punishment data (in shares) next to the estimates, so one can get a sense of what underlying data is driving the results. It should be clear from an examination of the graphs that they do not in any way change the conclusions we draw from the more aggregated versions.

In figures 10 through 13, the y-axis is the share of a recipient's total deduction points that were allocated to a DM with the indicated characteristics (i.e., the DM with a particular vote weight and proposal power). The hollow circles are point estimates of predicted punishment shares for a typical recipient (who allocated all 30 of her possible deduction points in a situation in which the DM's kept, a relatively selfish, 20 points of the initial endowment for themselves). These predicted effects are derived from the estimates in Table 3. The bars are 95% confidence bands calculated via simulation. The numbers to the right of each confidence band are the actual number of recipients allocating a given share of their deduction points to that DM (across all relevant experimental sessions) and so give a rough sense of how the underlying uncontrolled data drive the various predictions from the statistical model.

Figure 10: The Impact of Voting Weights and Proposal Powers on Punishment

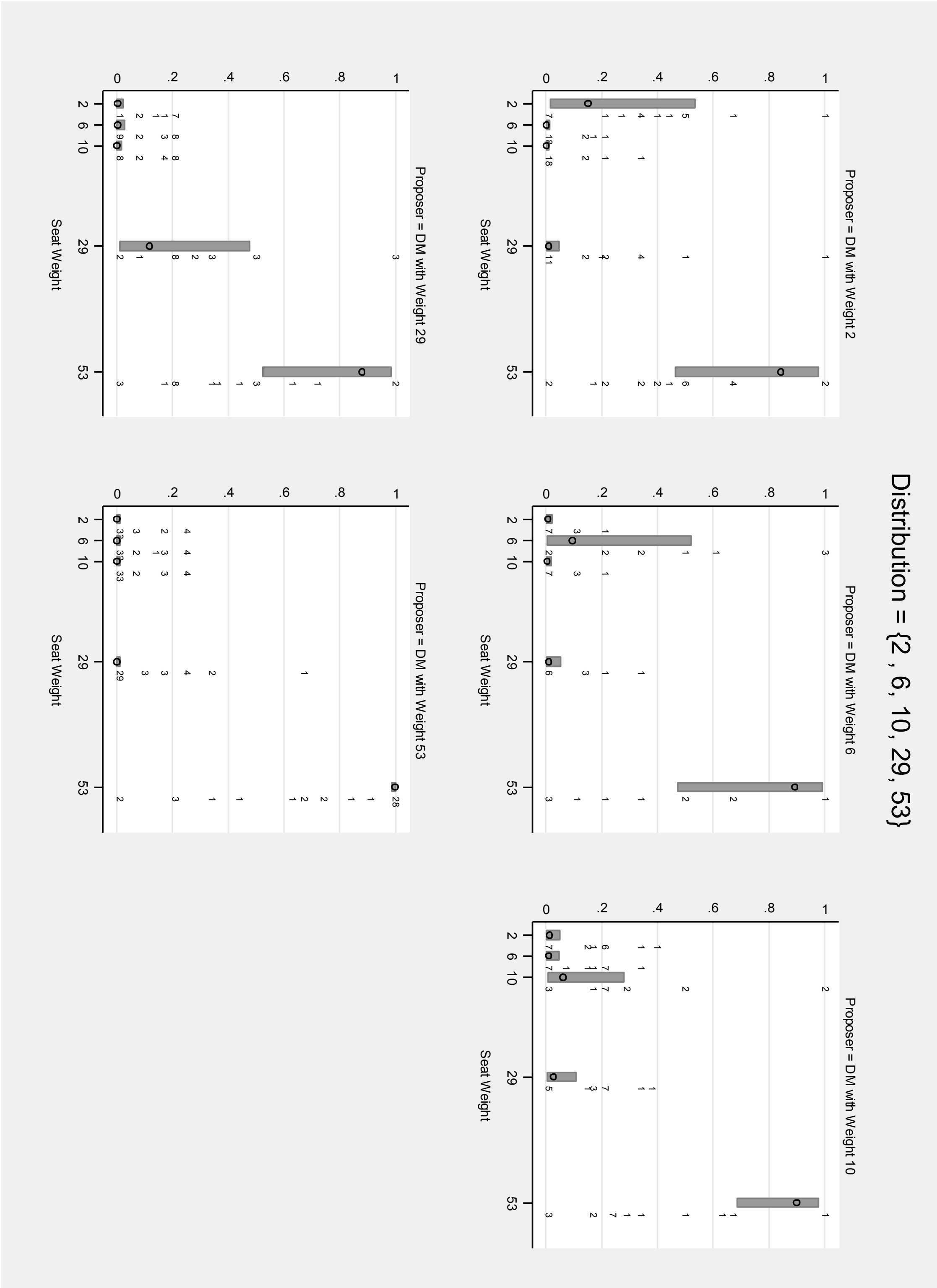


Figure 11: The Impact of Voting Weights and Proposal Powers on Punishment

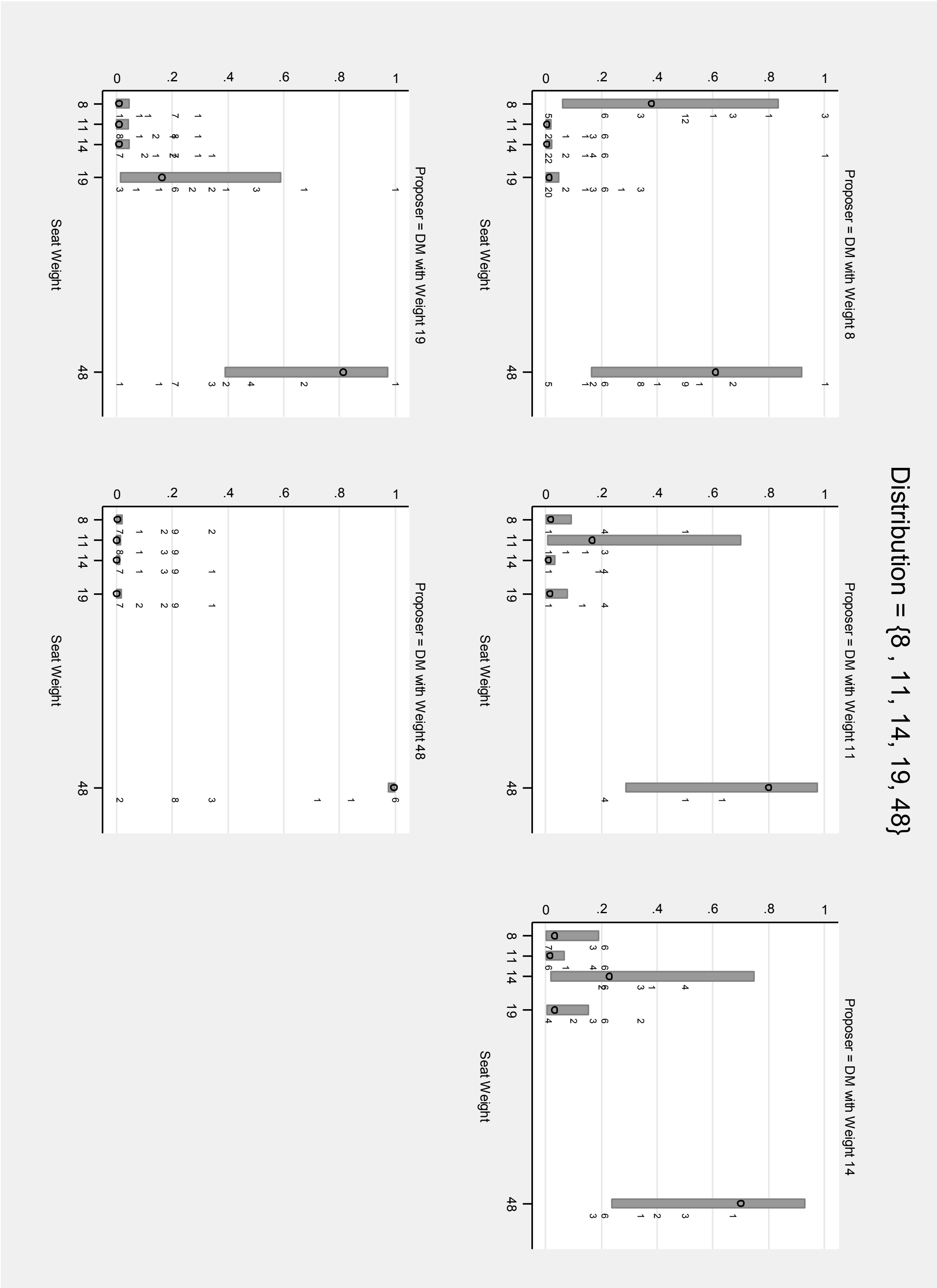


Figure 12: The Impact of Voting Weights and Proposal Powers on Punishment

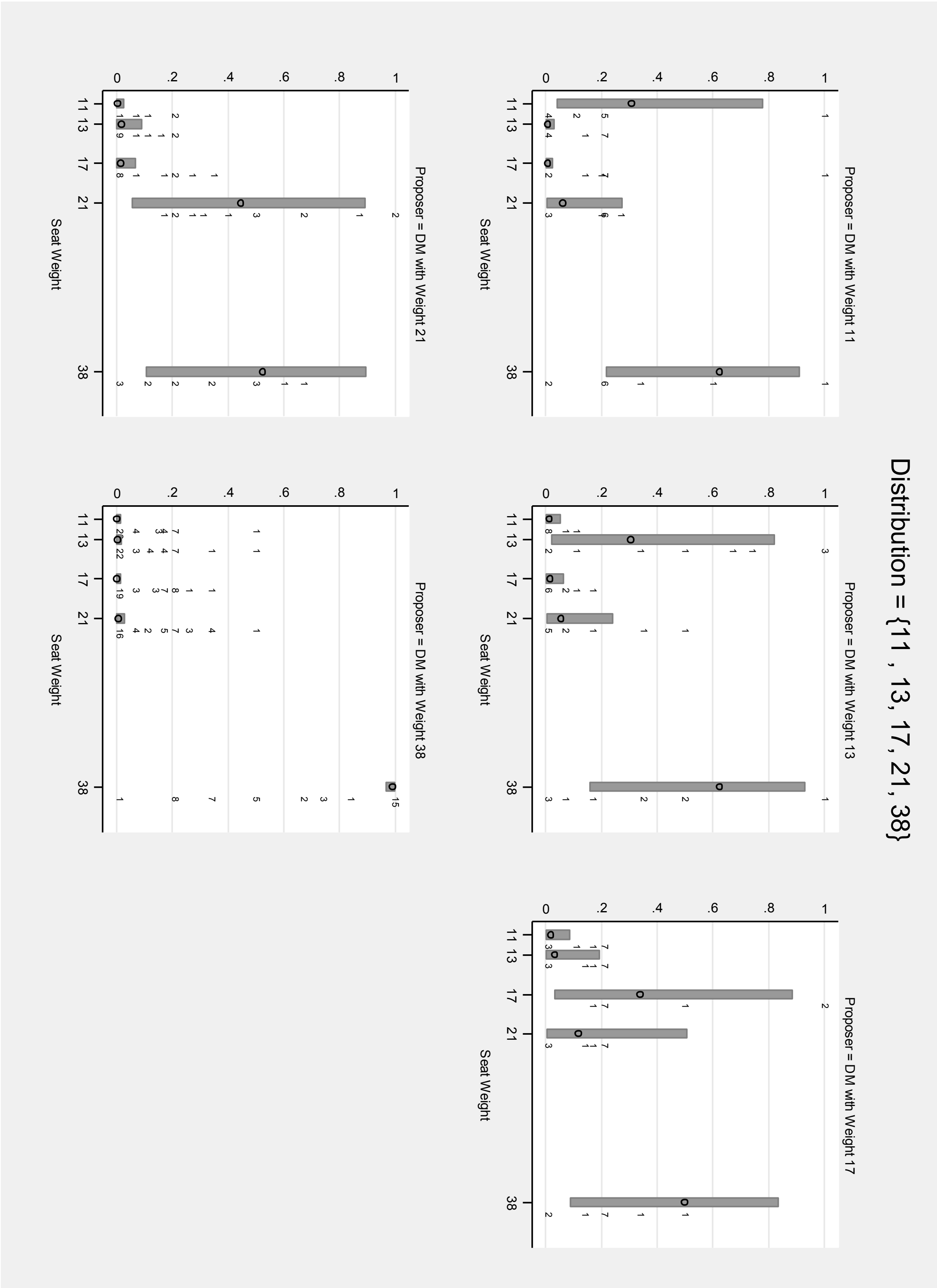
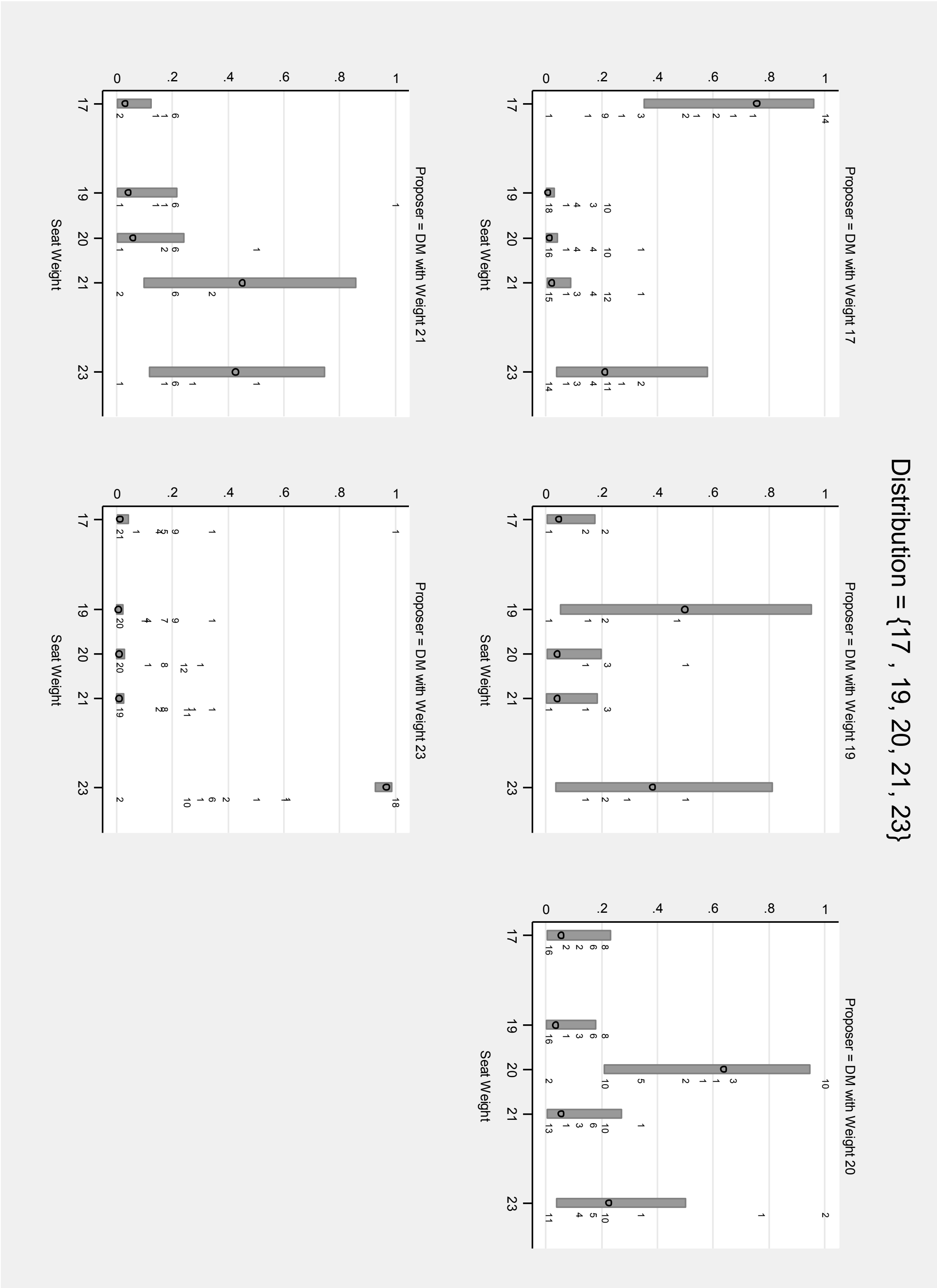


Figure 13: The Impact of Voting Weights and Proposal Powers on Punishment



3 Experiment 3: Internet Survey

3.1 Screen Shots

Figure 14: Internet Experiment Screen 1

This is a game in which you will be asked to guess the decision taken by five decision makers. We will describe three distinct decision making situations that took place earlier this week at the University of Oxford.

- In each situation, five individuals were given a total of £30 and asked to agree on how much should be given to two different charities, an animal shelter and a soup kitchen.
- Each has preferences for how much to give to the animal shelter (with the remainder going to the soup kitchen).
- The group as a whole, however, had to come up with a single, collective, amount to contribute to each charity.
- So, for example, if the group decided to donate £10 to the animal shelter then £20 were donated to the soup kitchen.

You are going to be given the following information:

- The contribution amount each of the five decision makers preferred (between £0 and £30) for the animal shelter.
- The rules they all used to come up with a collective choice.

At the end, we are going to ask you to guess what amount the five decision makers finally agreed upon.
We will now provide you with all of the information about the individuals that made the decision and the rules for coming up with a group decision about the amount to contribute to these charities.

Next

Figure 15: Internet Experiment Screen 2

The voting weights

- Each of the five decision makers was given a voting weight.
- The voting weights add to 1 and they determine how much each person's vote counts when the five people vote on the donation proposal.
- So for example, if there were five decision makers then the voting weights could be:
  - 0.4 for decision maker 1,
  - 0.1 for decision maker 2,
  - 0.2 for decision maker 3,
  - 0.2 for decision maker 4,
  - and 0.1 for decision maker 5.
- As you can see these voting weights add to 1.

Next

Figure 16: Internet Experiment Screen 3

The preferred donation

- Each of the five decision makers was randomly assigned a preferred donation amount for the animal shelter – it could be between £0 and £30
- For example, Decision Maker 1 could be assigned a preferred donation of 12 and Decision Maker 2 could be assigned a preferred donation of 21. This means that Decision Maker 1 prefers to donate £12 to the animal shelter and Decision Maker 2 prefers to donate £21 to the animal shelter.
- Each decision maker was rewarded with a monetary payoff that depended on whether the group of five decision makers could agree on a donation to the animal shelter that was similar to his or her preferred donation.
- For example, a decision maker whose preferred donation was 16 received his or her biggest monetary reward if the group choice was a donation of 16. And his or her monetary payoff got smaller as the group donation to the animal shelter got bigger or smaller than 16.

Next

Figure 17: Internet Experiment Screen 4

The voting process:

The exact process by which the decision making group voted a preferred allocation was as follows:

1. As mentioned before, the five decision makers were randomly assigned preferred donation positions and voting weights.
2. One of the five decision makers was randomly chosen to propose a donation amount (a number between £0 and £30) to be voted on by the group.
3. This proposal was then voted on by all five decision makers.
4. The voting rule requires a majority vote, as such a proposed collective choice can only be adopted if the sum of the voting weights of those in favour of the proposal is greater than 0.50, in which case the proposal was adopted.
5. If the proposal did not receive more than 50 percent of the weighted votes, then steps 3 - 5 were repeated until some proposal received more than 50 percent of the weighted votes and so was adopted as the collective choice.

Next

Figure 18: Internet Experiment Screen 5

Your role

- We are going to ask you to guess what amount the five decision makers finally agreed upon.
- The closer your guess is to the actual amount that was selected, the greater the number of SSI points you will receive.
- For instance, if you guess the exact amount agreed upon by the five decision makers, you will be awarded on average an additional 30 SSI points.
- However, the farther your guess is from the amount actually adopted, the fewer SSI points you will receive.
- If your guess is very far from the actual amount adopted, it is possible that you will receive 0 points.

Next

Figure 19: Internet Experiment Screen 6

Let's see an example...

Next

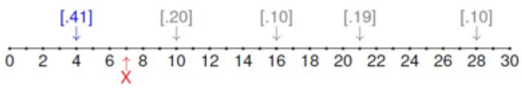
Figure 20: Internet Experiment Screen 7

In this example, the voting weights and the prefered donation positions of the five decision makers were the following (the decision maker who proposed the policy that was adopted is indicated in blue):

Decision Maker	Voting Weight	Donation Position
1	0.41	£4
2	0.20	£10
3	0.10	£16
4	0.19	£21
5	0.10	£28
TOTAL	1.00	

In this example Decision maker # 1 was the proposer.

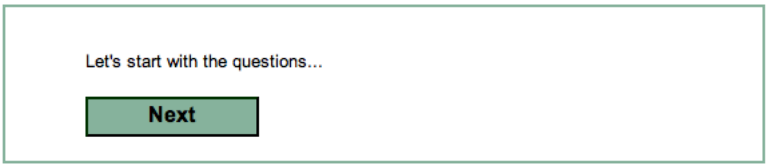
Below is a picture that shows the possible donation amounts (£0- £30), the location of the preferred donation amount of each decision maker (indicated by an arrow) and their voting weights (indicated in brackets). Again, the proposer's voting weight is indicated in blue



The outcome (which could be between £0 and £30) that got a majority vote in this made up example was £13.

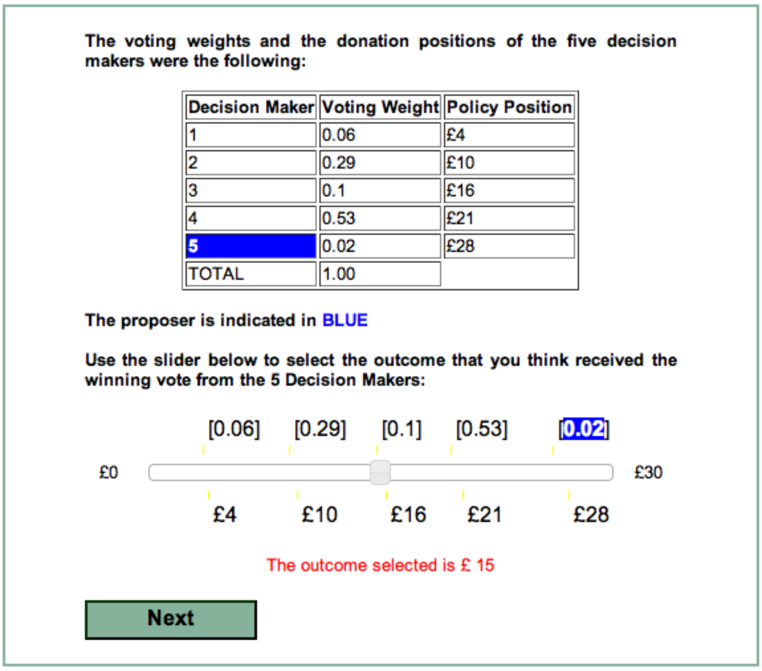
Next

Figure 21: Internet Experiment Screen 8


Let's start with the questions...

Next

Figure 22: Internet Experiment Screen 9


The voting weights and the donation positions of the five decision makers were the following:

Decision Maker	Voting Weight	Policy Position
1	0.06	£4
2	0.29	£10
3	0.1	£16
4	0.53	£21
5	0.02	£28
TOTAL	1.00	

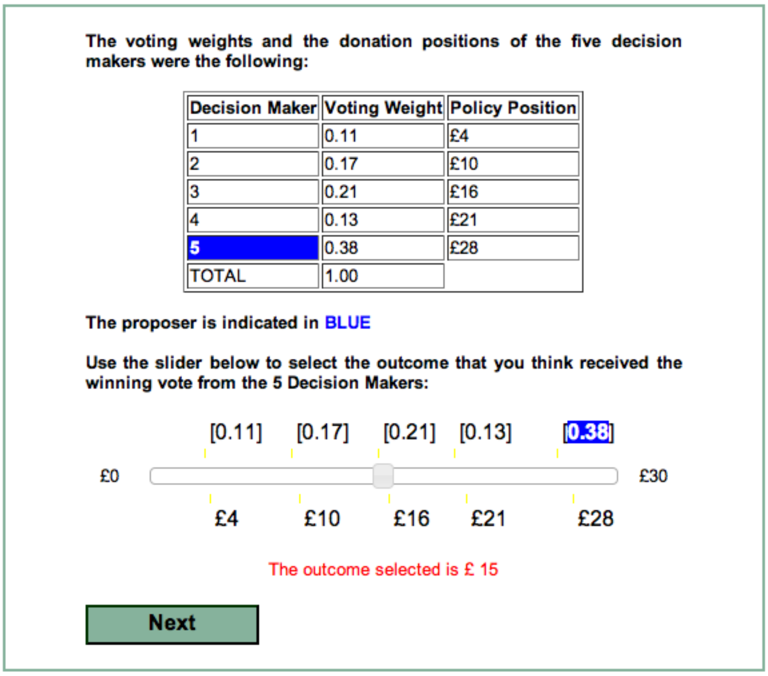
The proposer is indicated in BLUE

Use the slider below to select the outcome that you think received the winning vote from the 5 Decision Makers:

The outcome selected is £ 15

Next

Figure 23: Internet Experiment Screen 10

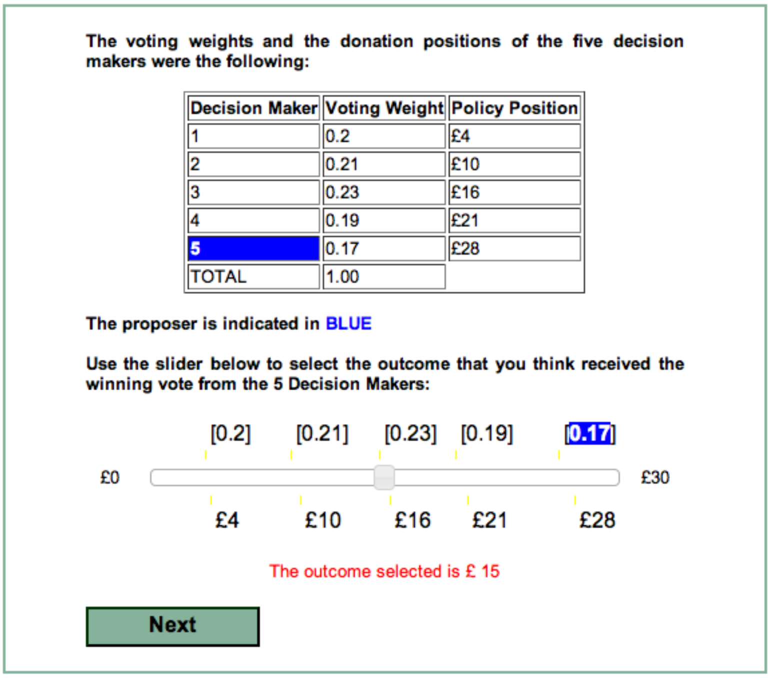

The voting weights and the donation positions of the five decision makers were the following:

Decision Maker	Voting Weight	Policy Position
1	0.11	£4
2	0.17	£10
3	0.21	£16
4	0.13	£21
5	0.38	£28
TOTAL	1.00	

The proposer is indicated in BLUE

Use the slider below to select the outcome that you think received the winning vote from the 5 Decision Makers:

Figure 24: Internet Experiment Screen 11


The voting weights and the donation positions of the five decision makers were the following:

Decision Maker	Voting Weight	Policy Position
1	0.2	£4
2	0.21	£10
3	0.23	£16
4	0.19	£21
5	0.17	£28
TOTAL	1.00	

The proposer is indicated in BLUE

Use the slider below to select the outcome that you think received the winning vote from the 5 Decision Makers:

3.2 Excel file defining random assignment to the five voting weights

An accompanying Excel file (appendix random assignment treatments.xls) contains the auxiliary values with which all of the randomisations were defined. Respondents are asked three questions concerning each of three voting weight distributions. There is a worksheet associated with each of these questions that describes how respondents were randomly assigned to voting weight treatments. *Treatment variation DSq1* corresponds to the majority distribution [.02, .06, .10, .29, .53]. *Treatment variation DSq2* corresponds to the unequal distribution [.11, .13, .17, .21, .38]. *Treatment variation DSq3* corresponds to the relatively equal distribution [.17, .19, .20 .21, .23]. There are 120 rows in each worksheet that correspond to the permutations of weights associated with each of the five Decision Making policy positions – in other words the weights can be ordered over the five decision making positions a total of 120 different ways. There are five columns in each spreadsheet corresponding to each DM position. Respondents were randomly assigned to a DM proposer treatment that correspond to these five columns. Each of the 600 cells in this spreadsheet corresponds to a randomly assigned treatment for the first distribution question – 120 permutations of the voting weights X 5 possible proposers. A similar random assignment for the other two questions is described in *Treatment variation DSq2* and *Treatment variation DSq3.*

The Excel file also contains three worksheets (*right location DSq1, right location DSq2, and right location DSq3*) that describe the "right" policy location for each voting weight and proposer randomization. These policy locations represent the policy outcome that was used to calculate the respondent's payoffs. The cells in these three worksheets correspond to the random assignment cells in worksheets *Treatment variation DSq1, Treatment variation DSq2, and Treatment variation DSq3.* And the distance between the respondent's guess regarding the policy outcome and this "right" policy outcome determined their payoffs. For example, the "right" policy position for a respondent randomly assigned to a row value of treatDSvar1=9 in *Treatment variation DSq3* with DM 1 would be 10.13. The distance between the respondent's guess and 10.13 would determine his or her payoff.

3.3 Multivariate Regression Results for Internet Experiment

In the main text we define the dependent variable for the results from the internet experiment as the spatial distance between the respondent's guess about the collective decision and the ideal points of each of the five DMs. We refer to this dependent variable in the regression model as distance. Table 6 reports estimates of three separate regressions (one for each question) of our distance variable on measures of the agenda powers, vote weights, and policy preferences of DMs. Specifically, we include an indicator variable for whether the DM was the proposer and expect a negative effect for this variable (i.e., respondents believe proposers are able to move collective policy choices towards their ideal point). We also include indicators for each vote weight in each distribution, leaving out the smallest weight as the baseline category. Thus, we expect (consistent with the results of the lab experiment) that having the largest vote weight will have a significant (negative) impact on distance, but expect no impact (or at least a much smaller and inconsistent impact) for vote weights more generally. Finally, we include a dummy variable for each DM's policy position (recall that the distribution of policy positions was the same for each question, though vote weights and agenda powers differed across questions and respondents).

Looking at the coefficients in Table 6, we see some initial evidence that respondents in our Internet survey experiment are acting much like the subjects in our lab experiments. This gives us some confidence that the cues we identified recipients use to apportion punishment are about identifying decision-making influence. Specifically, we see that the coefficient on the proposal dummy is strongly significant in each equation. Further, in the first and second equations, the dummy variable indicating the DM with the plurality seat weight (.53, .38, and .23 in the respective equations) is significantly different from both zero and the coefficients on the dummy variables for the sizes of other (smaller) DMs. In the last equation, no such effect is apparent. But given the very egalitarian distribution of vote weights used in this question (i.e., 17, .19, .20, .21, .23) this is really the exception that proves the rule: when size of the "largest" DMs is qualitatively not really different from the other DMs, then being the largest does not provide a very good cue for attributing influence.

Looking at the impact of vote weight more generally, the results are again largely consistent with our lab experiments. Indeed, the coefficients reveal that the impact of seat weight on distance (for DMs who are not the plurality DM) is often inconsistent (i.e., coefficients do not get monotonically larger in absolute terms for larger seat weights) and, with the partial exception of the first equation, insignificant (i.e., they are not all different from zero or from each other). Finally, the impact of policy position shows exactly the relationship we expected with distance being shortest for central DMs and increasing (symmetrically) as one moves to more extreme positions.

Table 6: OLS Regression of Individual Guesses and DM Characteristics

Variable	Distribution 1	Distribution 2	Distribution 3
Dummy for policy position 10	-5.43	-5.40	-5.17
	(0.07)	(0.08)	(0.08)
Dummy for policy position 16	-7.72	-7.95	-7.49
	(0.18)	(0.18)	(0.20)
Dummy for policy position 21	-5.32	-5.45	-5.17
	(0.26)	(0.26)	(0.28)
Dummy for policy position 28	0.80	0.79	1.08
	(0.31)	(0.30)	(0.32)
Dummy for vote weight .06, .13, .19	-0.43*[1]*	0.02*[2]*	0.17*[3]*
	(0.19)	(0.19)	(0.21)
Dummy for vote weight .10, .17, .2	-0.27	0.07	0.10
	(0.19)	(0.20)	(0.21)
Dummy for vote weight .29, .21, .21	-0.44	-0.33	-0.25
	(0.19)	(0.20)	(0.20)
Dummy for vote weight .53, .38., .23	-1.65	-1.02	-0.23
	(0.22)	(0.21)	(0.22)
Dummy for Proposer	-1.16	-0.99	-1.43
	(0.18)	(0.17)	(0.18)
Constant	12.44	12.09	11.82
	(0.22)	(0.19)	(0.22)
Rows of Data	5020	5020	5020
Respondents	1004	1004	1004
DMs per respondent	5	5	5

Notes: [1] Test to reject -.43=-.27=-.44=0 Pvalue=0.06; *[2]* Test to reject .02=.07=-.33=0 Pvalue=.17; *[3]* Test to reject .17=.10=-.25 Pvalue=.2. Reported standard errors are robust estimates clustered on individual respondent. We have also estimated hierarchical linear models that include error components (random effects) at the individual level. Results are quite robust to these alternative estimation strategies.