Essays in Experimental Political Economy

Jeffrey Da-Ren Guo

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2024

Abstract

Essays in Experimental Political Economy

Jeffrey Da-Ren Guo

In many economic applications, a collective outcome experienced by a group of people is determined by individual decisions made by its constituents. Hence, understanding how individuals make decisions in group settings is important, but empirical and observational analyses are often complicated by confounding factors. This dissertation contains three essays that use controlled experiments designed to isolate, and measure the impact of, mechanisms predicted to affect behavior.

Chapter 1 studies behavior under digital anonymity. A distinctive feature of the digital world is the ability to calibrate or withhold one's identifier: a person can be identified by a string of letters, an avatar, their real name, or even nothing at all. That digital identifiers allow a person to mask their physical identity also makes it difficult to attribute digital actions to a physical person, even when the actions are observed. I embed these features in an experiment where subjects play a finitely repeated, linear public goods game. Treated subjects are identified in one of three ways—by their photograph, by a random number, or by a self-designed cartoon avatar—and their individual choices are revealed and either attributed to, or decoupled from, their identifier. In line with the previous literature, identifying subjects and increasing the precision of attribution increases contributions relative to a baseline condition without identifiers or revealed individual choices. Remarkably, however, the largest impact on behavior comes from having an identifier in the first place: for a given level of attribution, the experimental data suggest that being

identified by a number or by an avatar is as powerful as being identified by one's photograph.

Chapter 2 studies whether and how individuals imbue digital avatars with self image and social image considerations. While digital avatars have become more commonplace and sophisticated, they need not resemble the physical appearance of the person using it. This inconsistency raises the question of how an avatar induces image considerations, relative to a person's physical appearance. I embed avatars into a dictator game and conduct two experiments, one addressing self image and the other social image. The direction of the treatment effect in the dictator game for both experiments suggests that individuals do attach image considerations to their avatars, though the effects are not statistically significant. Additionally, I find that subjects create significantly more positively perceived avatars when they know that their avatar will be shown to another subject who will decide how to allocate an endowment with them.

Chapter 3, joint with Alessandra Casella and Michelle Jiang, studies the impact of an alternative voting system on the minority's turnout and resultant victories. We start from the observation that under majoritarian election systems, securing participation and representation of minorities remains an open problem, made salient in the US by its history of voter suppression. One remedy recommended by the courts is Cumulative Voting (CV): each voter has as many votes as open positions and can cumulate votes on as few candidates as desired. Theory predicts that CV encourages the minority to overcome obstacles to voting: although each voter is treated equally, CV increases minority's turnout relative to the majority, and the minority's share of seats won. A lab experiment based on a costly voting design strongly supports both predictions. Chapter 3 was published in Volume 141 of *Games and Economic Behavior*, pp. 133-155, September 2023, `https://doi.org/10.1016/j.geb.2023.05.012`.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am indebted to Alessandra Casella and Mark Dean for their guidance and unwavering support. I thank Sharon Harrison, Navin Kartik, and Jacopo Perego for serving on my dissertation committee.

I thank my study group for hard carrying me through the first year: Michelle Jiang, Shin Oblander, Susie Scanlan, and Qianyang Zhang.

I thank my co-authors, through whom I experienced firsthand the enjoyment that can be had from doing research: Pietro Battiston, Alessandra Casella, Mark Dean, Sharon Harrison, Hayeon Jeong, Zhi Hao Lim, and Michelle Jiang.

I thank the Columbia Experimental Laboratory for the Social Sciences for its wealth of resources: financial support for experiments and travel, the robust subject pool, the physical lab space, the Experimental Lunch seminar, and the SWEET conference. I also thank the Lab Managers who have been a huge help to me in testing and conducting my experiments: Silvio Ravaioli, Victoria Mooers, and Shuhua Si.

I thank the Program for Economic Research for their financial support.

I thank Victoria Geh for her research assistance in coding the experiments in Chapter 2.

I thank Amy Devine for her indispensible work as the PhD Program Coordinator and the Graduate Program Manager. I thank Navin Kartik for his tireless work as the Director of Graduate Studies.

I thank the Association of Graduate Economics Students for its efforts in advocating for, and its successes in creating a community among, PhD students.

Dedication

To my parents, without whom I would not have started.

To Abigail, without whom I would not have finished.

Chapter 1: Doing Good in the Digital World

1.1 Introduction

The internet has given rise to digital spaces in which individuals gather and interact, from AOL chat rooms in the 1990s to modern-day forums like Reddit that attract over 50 million daily active users.[1] Technological advancements have facilitated digital interactions that more closely resemble those of the physical world around us. Massively multiplayer video games like *World of Warcraft* attract millions of players who walk around to meet each other's characters and complete tasks together. Platforms like *Second Life* and *VRChat* bring together thousands of users simply to hang out with each other, but as avatars in virtual worlds. And in recent years, developments in augmented and virtual reality technology have brought the physical and digital worlds closer together than ever before. As the line between the physical and digital continues to blur, it has become imperative to understand how individuals behave in the digital world, particularly because myriad anecdotes of online toxicity, harassment, and abuse remain pervasive.[2] How can the digital world be designed so as to induce people to do good unto one another?

One key distinction between interactions in the digital and physical worlds is how participants are identified to one another. In most physical interactions—running into someone on the street, gathering with friends for dinner—each participant is identified by their literal, physical person. Simply by showing up to an interaction in the physical world, one inherently brings along an identifier that uniquely pins down, or identifies, who they are. On the other hand, identifiers in the digital world can take many forms and vary in how precisely they identify the underlying person. In digital interactions, one could be identified by a random string of letters and numbers (imprecise),

[1] `https://www.redditinc.com/blog/reddits-2020-year-in-review/`

[2] A 2021 *New York Times* article opens with the headline, "The Metaverse's Dark Side: Here Come Harassment and Assaults." In the same year, the Pew Research Center released a report that found that 41% of Americans had experienced some form of online harassment.

by a picture of their face (very precise), or even by nothing at all. And in recent years, lifelike avatars have become an increasingly common digital identifier, further blurring the line between the digital and physical.[3]

An inextricable consequence of the wide range of digital identifiers is that attributing actions to the underlying actor becomes more difficult. In most physical interactions, when an action is observed, it is naturally attributed to the person who performed it. Past studies find that unmasking physical identities and attributing actions have a complementary, positive effect on the level of prosocial behavior ([3]; [4]). But, it is unclear how this effect translates to digital settings, where actions need not be attributed to an identifier. Moreover, even if actions can be attributed to a digital identifier, the identifier itself may not identify the person behind the keyboard.[4] In short, digital settings allow an easier disconnect between a person and their identifier, and between a person and their actions. This raises the main questions that the paper seeks to answer: *Can less precise identifiers and less precisely attributed actions still induce prosocial behavior? Which channel is more powerful in inducing prosocial behavior?*

Thus, I conduct an experiment whose two arms of variations are in identification and the attribution of individual actions. First, subjects have identifiers that differ in the precision with which they identify the physical person. Treated subjects are identified to each other by either their photograph, a self-designed cartoon avatar, or a random three-digit number; *Baseline* (control) subjects are not identified to each other at all. The second arm of variation is in the attribution scheme used to publicly attribute subjects' choices in the game to their identifiers. Treated subjects' individual choices are publicly revealed and either attributed to (*Full Attribution*), or decoupled from (*Partial Attribution*), their identifier; *Baseline* subjects' individual choices are kept private.

I embed this variation in a linear public goods game ([5]) played in fixed groups of four subjects each, finitely repeated for thirty-two rounds. Public goods games have been studied extensively

[3]The realism of avatars lends credence to the popular argument that an avatar is an extension of one's person. For example, Meta writes on their website at `https://www.meta.com/avatars/` that "[a]vatars are a digital expression of you, letting you freely express your identity, personality and appearance."

[4]Interactions on Reddit are an example of this. Each Reddit user's history of posts and comments are logged under their username. However, Reddit does not require its users to use their real name (orthonym) as their username.

in the laboratory to address questions about prosocial behavior in group settings.[5] My experiment implements a $(3 \times 2) + 1$, between-groups design. The *Baseline* treatment replicates the canonical implementation of the public goods game in the lab: subjects are not identified to each other, and only the aggregate contribution to the public good is revealed after each round.[6] In each of the treatment conditions, two elements are added to the game. First, each subject has an identifier: *Number*, *Avatar*, or *Photo*. Second, each subject's contribution choice is also revealed to the group after each round—the breakdown of the aggregate contribution is made known. When the *Full Attribution* scheme is used to reveal choices, each subject's contribution choice appears underneath the subject's identifier. In contrast, when the *Partial Attribution* scheme is used, subject contributions are presented in descending order, from largest to smallest, without being linked to any specific identifier. Each of the six treatment conditions is characterized by the Identifier and the Attribution scheme shared by all members in a group. This dictates how I abbreviate the treatment conditions: for example, I will abbreviate the treatment condition with *Photo* and *Full Attribution* as *PhotoFull*.

Strictly speaking, neither adding identifiers nor revealing individual choices changes the subgame perfect Nash equilibrium prediction of the game, which states that players contribute zero ("free ride") in all periods of the game.[7] However, I analyze two standard channels that may induce positive contributions. The first is image concerns.[8] In a simple theoretical framework, I derive two predictions on equilibrium contributions when agents also derive utility from their image. These predictions revolve around two parameters. One parameter (b) captures the intensity with which an agent cares about, identifies with, or invests in their identifier[9]; the other parameter (p) cap-

[5] [6] and [7] survey results from public goods experiments in the laboratory.

[6] See, for example, the experimental implementation of the public goods game in [8].

[7] It is well known that experimental subjects do not play the Nash equilibrium of contributing zero: rather, subjects contribute some (but not all) of their endowment in early rounds, and contribution levels deteriorate over time ([6]; [9]).

[8] Impure altruism, or warm glow, is a commonly cited source of self image utility ([10], [11]). Various social effects—prestige, recognition, shame—are sources of utility derived from one's social image ([12]; [13]).

[9] There is evidence that one's identifier can affect the level of self-image concerns one faces. [14] provides experimental evidence that a decisionmaker faces higher levels of self-image concerns when they see themselves in a mirror when making their decision, compared to seeing nothing or a neutral stimulus.

tures the probability with which the agent's action can be attributed to their identifier.[10] The first prediction is that equilibrium contributions are increasing in b and p. Utility-maximizing agents contribute more when their contribution is more likely to be attributed to their identifier. Taken to the experiment, this predicts higher contributions under *Full Attribution* than under *Partial Attribution*. Additionally, assuming that subjects identify less more with their photograph than with a number (*i.e.*, $b_{\text{Photo}} > b_{\text{Number}}$), this also predicts higher contributions under *Photo* than under *Number*. The second prediction is that contributions have increasing differences in (b, p): the impact of a change in Identifier should be larger under *Full Attribution* than under *Partial Attribution*, and the impact of a change in Attribution should be larger under *Photo* than under *Number*. This follows from b and p being complements, as is consistent with the literature.

The second channel is the possibility that the game extends outside the laboratory. Imagine that group members are able to sanction each other *outside* the laboratory, even after the public goods game in the laboratory is over.[11] Could the possibility of being punished outside the public goods game induce positive contributions during it? I consider an infinite horizon supergame that consists of two games: a finitely repeated public goods game, followed by an infinitely repeated costly punishment game.[12] I show that the threat of punishment can indeed induce positive contributions in the public goods game, but that the size of the threat depends on whether group members can identify and coordinate punishment on a free rider. I show that the expected punishment from free riding is strictly higher when contribution choices are (1) attributed with certainty to an identifier that (2) can identify the free rider outside the public goods game. This result, when applied to the experiment, predicts higher contributions in the *PhotoFull* condition than in all other conditions.[13]

[10] In [15]'s model of decisionmaking with image concerns, the reputational payoff of taking an action is increasing in the "visibility" of the action—the "probability that it will be observed by others."

[11] Many studies consider the effect of being able to punish free-riders *within* the public goods game itself by incorporating a punishment scheme into the game. In such schemes, subjects choose if they want to punish other members of their group in between rounds of the public goods game (after observing the contribution choices of all group members), usually at a cost. [16] implement a similar sanctioning scheme within a common pool resource game. The addition of a punishment scheme has been found to be an effective mechanism for inducing prosocial behavior ([9]; [17]; [18]; [19]; [20]).

[12] Imagine a group of friends which gathers one evening to play a board game. Actions taken in the game may spill over into their larger friendship: a person who behaves selfishly in the game may be shunned by the other friends afterward.

[13] Note that what I model here is different from an infinitely repeated public goods game. Analysis of an infinitely

Relative to the *Baseline* condition with neither identifiers nor revealed individual choices, I find more prosocial behavior—higher contributions—in all six treatment conditions. The increases are statistically significant at the 5% level for five of the six treatment conditions; moving from *Baseline* to *PhotoFull* increases contributions by 68%. However, changes in Identifier and Attribution do not affect behavior equally. The main finding of the paper is that Attribution is the more powerful channel in inducing prosocial behavior. While average contributions are indistinguishable across the three different Identifiers (pooling across Attribution schemes), there is a significant increase in contributions from *Baseline* to *Partial Attribution*, and again from *Partial Attribution* to *Full Information* (pooling across Identifiers). In fact, under *Full Attribution*, contributions are just as high when subjects are identified by a random number or a cartoon avatar as when they are identified by their photograph. Put another way, the data suggest that mere presence of an Identifier is sufficient to induce higher contributions, as long as the Identifier is paired with information about individual choices.

These results are interesting in light of the experiment in [3], whose 2×2 design incorporates variation on the same two dimensions: whether subjects are identified, and whether their individual contribution choices are revealed. Their experiment has a condition in which subjects are not identified and only the aggregate contributions are revealed (analogous to my *Baseline* condition), as well as a condition where subjects are identified by their photograph and their individual contributions are attributed to their photo (analogous to my *PhotoFull* condition). In their two intermediate conditions, one of the two arms is entirely shut off: in one condition, subjects are identified by their photographs but only aggregate contributions are revealed; in the other condition, subjects have no identifiers, and individual contribution choices are displayed from largest to smallest. A significant increase in contributions (relative to their Baseline) comes only when identification with

repeated public goods game using the Folk Theorem suggests that positive contributions can be sustained in equilibrium for sufficiently large discount rate. Notably, there exist common "trigger" strategies, such as Grim Trigger, that can sustain positive contributions in equilibrium with only feedback about the *aggregate* contribution between rounds—neither identifiers nor revealing individual contribution choices is necessary. If subjects mistakenly believe the public goods game in the laboratory to be infinitely repeated, and play such trigger strategies *as if* the game were infinitely repeated, then we would expect no difference in behavior between the *Baseline* condition and the treatment conditions.

photographs is combined with revealing individual contribution choices. As they write, "adding just information on generosity has no significant effect on giving, and neither does adding just the identity of the giver. However, a substantial impact comes from using both in combination." The treatment conditions in my experiment maintain the presence of both identifiers and individual contribution choices, and instead vary the precision of the identifiers and the precision of the attribution of the choices to identifiers. Indeed, in five of my six treatment conditions, contributions are significantly higher than in the *Baseline*: this suggests that Andreoni and Petrie's result is robust to settings where identifiers and attribution are imprecise, but present.

Additionally, the experimental data are inconsistent with the punishment channel, and not fully consistent with the image channel. As a result, I conduct exploratory analysis to identify what is driving contribution behavior. I find that the increase in contributions from *Baseline* to *Partial Attribution* to *Full Attribution* is driven primarily by an increase in the rate of subjects contributing their full endowment: the rate of full contribution rises from 16.6% in *Baseline* to 38.6% in *Partial Attribution* to 62.3% in *Full Attribution*. Note that, when individual contribution choices are revealed, subjects are also given pieces of "relative" information: for example, the rank of their contribution in the group, the minimum and maximum contributions, whether an other member of the group contributed their full endowment or contributed nothing.[14] To test whether relative information influences future contribution behavior, I estimate a dynamic panel model with the experimental data using the Arellano-Bond estimator [21]. In particular, I estimate the model separately for *Baseline*, *Partial Attribution*, and *Full Attribution*: not only do I check that a piece of relative information has a significant effect on future contributions under *Partial Attribution* and *Full Attribution*, I also check that it has no significant effect in the *Baseline*.

Indeed, the dynamic panel estimation shows that subjects condition future contributions on "relative" information gleaned from past rounds, but only when it is available. Under *Partial Attribution* and *Full Attribution*, subjects significantly reduce their contributions in period t when

[14] "Aggregate" information, on the other hand, can be known from having only the aggregate contribution to the group account: the total amount contributed by other group members, whether one's contribution was above or below the mean contributed by other group members.

another group member free rides in period $t-1$; this does not happen in the *Baseline*. Moreover, the amount by which subjects increase their contributions from period $t-1$ to period t is increasing in the number of group members that contributed strictly more than the subject in period $t-1$—but only under *Partial Attribution* and *Full Attribution*. Taken together, these results suggest that when individual contribution choices are available, subjects use the additional information to establish a norm of reciprocity for their group.[15] When a subject knows that they contributed a relatively low amount, they increase their contribution in the next period; yet, when a subject realizes someone else violated the norm by free riding, they reduce their contribution in the next period. Interestingly, the analysis shows that the effects of reciprocity are stronger under *Full Attribution* than under *Partial Attribution*, even though the amount of relative information is the same under both Attribution schemes. This suggests that not just reciprocity, but also an interaction between reciprocity and image, drives behavior.

The experimental data also shed light on how individuals choose to identify and represent themselves in the digital world. From the *Avatar* condition, I observe the cartoon avatars created by 100 subjects (56 assigned to *AvatarFull*, 44 assigned to *AvatarPartial*), as well as the amount of time each subject spent on customizing their avatar. Subjects were able to choose visual attributes from between twelve dropdown menus; in total, over 2.8 trillion permutations of attributes were available to be chosen.[16] Additionally, after the public goods game ends, *Avatar* subjects reported their race and gender, as well as the intensity with which they felt represented by their avatar.[17] I find that subjects invested substantial time in customizing their avatar: on average, subjects spent just over three minutes making their avatars.[18] However, I do not find a correlation between a subject's customization time and their contribution in the first round of the game. I do find a

[15] [22] also finds contribution behavior in public goods games that is consistent with reciprocity models (over commitment or altruism models). In particular, they find that subjects in round t try to match the median contribution from round $t-1$, as opposed to the minimum or maximum contribution. In a field experiment, [23] find evidence of reciprocity in giving. For a general discussion of the theory of reciprocity, see [24].

[16] The cartoon avatars looked like passport or ID photographs: a "face" from the "shoulders" up. Figure 1.1 shows the dropdown menus available, as well as a sample avatar.

[17] Subjects rated their agreement on a 7 point Likert scale (1 being Strongly Disagree, 7 being Strongly Agree) with the following statement: *I felt like my avatar represented me.*

[18] I find no significant difference in customization time between *AvatarPartial* and *AvatarFull* subjects.

positive, albeit weak, correlation between self-reported representation and first round contribution among *AvatarPartial* subjects (but no such correlation for *AvatarFull* subjects). Finally, there is evidence that some subjects created avatars that look different from their physical person, based on their reported race and gender.

1.1.1 Literature

This paper contributes primarily to a rich literature that studies the effect on prosocial behavior of identification and attribution of individual actions. Much experimental work finds a positive effect of both channels on prosocial behavior in both lab and field settings.[19] [26] identify some players in dictator games by their family names, and find that dictators allocate more of the endowment to the other player when names are known. [27] find that subjects exert more effort in a real-effort task for charity when they are compelled to reveal the amount of money they earned to other participants in the lab. In the field, [28] studies charitable giving at churches in Denmark by manipulating the visibility of others' donations in the collection vessel; churchgoers give more when the collection vessel allows others' contributions to be seen. [29] pair a field experiment conducted through an on-campus charitable organization with a lab experiment that reveals donation decisions of some subjects to others in the room. Evidence from their experiment suggests that individuals give in order to improve their social image, rather than purely altruistically. Indeed, the results of my experiment are consistent with this pattern: subjects behave more prosocially when they are identified, and when their actions can be attributed back to them.

A similar effect has been found in public goods games: identification and attribution have a positive effect on contributions. In the lab, variation in attribution was first implemented by revealing individual contribution choices in addition to the aggregate contribution. Results from past studies are mixed: [30] find a significant increase in contributions in the last five rounds of their experiment, while [31], [32], and [33] do not find an effect of revealing individual choices on

[19] [25] do not find a positive effect of identification and attribution in their lab experiment. However, they note that their experimental implementation introduces confounding factors.

contributions.[20] In a related study, [34] find that presenting "relative" information to subjects—including the maximum and minimum contribution choices and the ranking of one's contribution in the group—increases contributions. More recently, experimenters have also identified the players themselves, along with revealing their choices. [4] find an increase in contributions when players are compelled to count out their contribution and write the amount on a blackboard in front of the group. [13] identify players using their photographs and first names: they find that revealing the two lowest contributors and their choices each round induces the same increase in contributions as revealing all players and their contributions. On the other hand, [3]'s treatment with photographs, but only aggregate contributions, doesn't find a statistically significant increase in contributions compared to their Baseline. Moreover, their treatment with individual contribution choices, but without identifiers, finds a decrease in contributions relative to their Baseline. The results from my experiment suggest that identifiers and individual choices are needed in conjunction to increase contributions, as is consistent with previous findings. The novel finding from my experiment is that this result still holds, even when both the identifier and the attribution scheme are imprecise.

Also related is a smaller literature that studies the relationship between one's avatar and behavior. [35] establish a physiological effect of being able to choose one's avatar: in their experiment, subjects who customized an avatar had a 10% faster heart rate compared to those who were assigned their avatar. Other studies have shown an effect of a person's avatar on their behavior. [36] find that subjects who were assigned tall avatars in virtual reality behaved more confidently than subjects assigned short avatars.[21] The effect of an avatar has also been shown to extend into the physical world: [37] find that subjects who were assigned a black avatar in virtual reality subsequently exhibited greater racial bias in the physical world than those assigned a white avatar. Moreover, avatars have been shown to be effective substitutes for photographs in reducing social

[20] It should be noted that in the [30] and [32] experiments, individual contribution choices were presented in a way that allowed a subject's contribution history to be tracked over time. In contrast, [33] presents individual contribution choices in ascending order. [31] does not describe how individual contribution choices were presented, but does note that "[s]ubjects have no contact before, after and during the game: they act in strict anonymity."

[21] In their experiment, subjects wore virtual reality 'helmets' and played multiple rounds of an ultimatum game against a confederate. Subjects assigned tall avatars proposed splits that were significantly more in their own favor than subjects assigned short avatars. Moreover, subjects assigned tall avatars were half as likely to accept an unfair offer from a confederate as subjects assigned short avatars.

distance, and increasing cooperation and trust between individuals.[22] On the other hand, there is evidence that people don't simply create avatars that look like themselves. [41] and [42] find that users in virtual worlds consider the types of people they will interact with as well as the nature of the interactions themselves when customizing their own avatars.[23] Recent experimental work in economics has found that strategic considerations could also drive such behavior. [43] find that, when competing to be "hired" for a math task, female subjects choose female avatars significantly less often than when the nature of the task is unknown. [44] also find a gender effect in a marketplace experiment: they find that buyers trust female avatars more than male avatars. In response, sellers with "genuine" male avatars are more likely to switch to female avatars than vice versa.

Compared to past studies, my experiment gives subjects many more degrees of freedom in creating their avatars: while subjects in [43] chose one of three avatars and subjects in [44] chose one of eight avatars, subjects in my experiment had billions of avatars available to them via the dropdown menus. Given this freedom, I find that subjects do invest time in customizing their avatars, though the time spent customizing does not depend on whether actions will be attributed to their avatars with certainty or only probabilistically. Furthermore, I find some evidence that subjects don't make avatars that look like their physical person, though my design does not allow me to identify the reason.

The paper proceeds as follows. Section 1.2 describes the experiment. Section 1.3 discusses the theoretical framework. Section 1.4 presents the results, and Section 1.5 concludes. The formal theoretical framework is left to the Appendix. Experimental instructions and screenshots appear in the Appendix.

1.2 Experimental Design

The game in the experiment is the finitely repeated, linear public goods game (also known in the literature as the voluntary contribution mechanism). I begin by describing the one-shot stage

[22] See, for example, [38], [39], and [40].

[23] [41] study subjects who were asked to create an avatar on the Yahoo! Avatars site either for blogging, dating, or gaming. [42] analyze observational data, survey responses, and interviews of users in the *Second Life* virtual world.

game and the finitely repeated game, as well as the standard theoretical predictions for those games. Next, I describe the experimental design and implementation.

1.2.1 The Linear Public Goods Game

The stage game is parameterized by $\{N, m, Y\}$, where N agents are in a group together, and each is endowed with Y points. Each agent $i \in \{1, ..., N\}$ privately chooses an amount of points $g_i \in [0, Y]$ to contribute to a group account (the public good), and keeps the remaining $Y - g_i$ points in a private account. Every point contributed toward the group account is multiplied by a factor m, where $m > 1$ but $m < N$. Total contributions to the public good are redistributed to the agents in equal shares. Hence, agent i's payoff in the stage game is given by:

$$\pi_i(g_1, ..., g_N) = (Y - g_i) + \frac{m}{N}(g_i + \sum_{j \neq i} g_j) \tag{1.1}$$

Each agent's marginal per capita return (MPCR) from contributing a point to the group account is $\frac{m}{N} < 1$, while the MPCR from keeping a point in their private account is 1. Assuming an agent seeks to maximize only their own payoff, the agent has a dominant strategy of contributing zero to the group account in the stage game: I call such a strategy *free riding*. Therefore, all agents free ride in the unique Nash equilibrium of the stage game: $(g_1, ..., g_N) = (0, ..., 0)$; the equilibrium payoff for all agents is equal to Y—the initial endowment.[24]

The stage game can also be repeated. Suppose the public goods game is repeated finitely for $T > 1$ discrete periods indexed by t. Denote agent i's contribution in period t as g_{it}. I assume that the aggregate contribution to the group account in period t is made known to all group members, before contribution decisions are made in the next period $t + 1$. Formally, denote the aggregate contribution in period t as $G_t \equiv \sum_{i=1}^{N} g_{it}$. Thus, for all periods $t > 1$, the history of past aggregate contributions $\boldsymbol{G}_t \equiv \{G_1, ..., G_{t-1}\}$ is common knowledge.

Standard analysis of the repeated game is straightforward. Because the game is finitely re-

[24]Note that this outcome is not Pareto optimal because $m > 1$: each point becomes bigger when contributed to the group account. The unique Pareto optimal outcome in the stage game occurs when all agents contribute their entire endowment to the group account—that is, $(g_1, ..., g_N) = (Y, ..., Y)$, and all agents earn a payoff equal to mY.

peated, and it is common knowledge that the game is finitely repeated, the game in the final period T can be analyzed in the same manner as the stage game: the unique Nash equilibrium in the final period is $(g_1, ..., g_N) = (0, ..., 0)$. Backward induction then yields that the unique subgame perfect Nash equilibrium remains: $(g_1, ..., g_N) = (0, ..., 0)$ in each round.

Note that if the game were instead infinitely repeated, then the Folk Theorem would apply: positive contributions could be sustained in equilibrium. In particular, note that if the game were infinitely repeated, it is possible to sustain positive contributions while knowing only the level of *aggregate* contributions in previous periods.[25]

1.2.2 Experimental Design

In all sessions, subjects played the finitely repeated, linear public goods game just described in fixed groups. Each subject participated in exactly one session and belonged to exactly one group: subjects stayed in the same group for the duration of the experiment. The game was parameterized as follows: groups were of size $N = 4$ and the game was finitely repeated for $T = 32$ rounds. In each round, each subject was endowed with $Y = 20$ points and privately chose an integer amount $g \in \{0, ..., 20\}$ to contribute to the group account; the remaining $20 - g$ points would be kept in their private account. All contributions to the group account were doubled ($m = 2$), so that $\frac{m}{N} = 0.5$. After each round, subjects were told the aggregate contribution to the group account (G), as well as their own earnings from that round. At the end of the experimental session, subjects were paid in cash based on the total number of points accumulated across all 32 rounds, converted at a rate of 50 points to \$1.00 (or 1 point to \$0.02).[26]

[25] Per the Folk Theorem, it is possible to sustain strictly positive contributions in equilibrium as long as agents are sufficiently patient (*i.e.*, the discount factor $\delta \in (0, 1)$ is sufficiently large). In context of the public goods game, it has been shown that such an equilibrium can be sustained by many different types of "trigger" strategies, which condition future contribution(s) on the history of past aggregate contributions ($\boldsymbol{G_t}$). Such trigger strategies take the following general form: *Contribute a strictly positive amount $g > 0$. If an other player deviates and contributes less than g, then punish the deviation by contributing $g = 0$ in the next period. Continue to contribute $g = 0$ for some number of periods, after which return to contributing the strictly positive amount $g > 0$.* A common trigger strategy is Grim Trigger, which says to contribute one's entire endowment ($g = Y$) until an other player deviates, at which point one should contribute zero forever. But it can be shown that other trigger strategies, which dictate a different amount to contribute g and/or a different duration of punishment, can be supported in equilibrium as well ([45]).

[26] Subjects were paid one at a time, in private. Subjects saw only their own earnings, and were instructed not to discuss their earnings with others.

The experiment used a between-groups, $(3 \times 2) + 1$ design. The *Baseline* condition implements exactly the finitely repeated game as described above: only the aggregate contribution to the public good is made public after each round, and subjects are not identified to each other in any way.[27] The six treatment conditions introduce two additional elements to the game. First, each subject is identified by an identifier in the game: either a randomly selected three-digit number (*Number*)[28], a photograph taken of the subject (*Photo*)[29], or a cartoon avatar customized by the subject (*Avatar*).[30] I abbreviate each treatment condition in two words: one for the Identifier, and one for the Attribution scheme. For example, I write the treatment condition with *Photo* and *Full Attribution* as *PhotoFull*.

In *Avatar* sessions, subjects were given five minutes before the start of the game (but after instructions were read) to customize their avatar by selecting different visual attributes from drop-down menus. The avatar customization screen was coded in a way that allowed subjects to see, in real time, how changing a particular attribute would affect the appearance of the avatar. Hence, subjects could experiment with different attributes—and see how they would look—before finalizing their avatar. Figure 1.1 shows screenshots of the avatar customization interface.

Second, subjects are *additionally* given, after each round, the individual contribution decisions made by each group member. These individual contribution decisions are either attributed to the identifier of the subject that made the contribution (*Full Attribution*), or decoupled from the identifiers (*Partial Attribution*). *Partial Attribution* of individual contribution decisions was implemented by presenting the contribution amounts from highest to lowest after each round, while the order of identifiers remained constant throughout the session. Figure 1.2 shows how *Full Attribution* and *Partial Attribution* were implemented in the experimental interface.

[27]The Baseline condition also corresponds to the canonical implementation of the public goods game in experiments.

[28]In *Number* sessions, no two subjects in the same group had the same number.

[29]In *Photo* sessions, each group member was identified by photograph taken at the start of the experimental session. The photographs were taken one person at a time by the experimenter, using a digital camera. All photographs were taken in public view of all participating subjects.

[30]The universe of cartoon avatars used in the experiment are from an open source Sketch library designed by Pablo Stanley, available at `https://www.avataaars.com/`. The avatar customization interface and dropdown menus used in the experiment were based on, and supported by, an open source web-based app developed by Fang-Pen Lin, available at `http://getavataaars.com`.

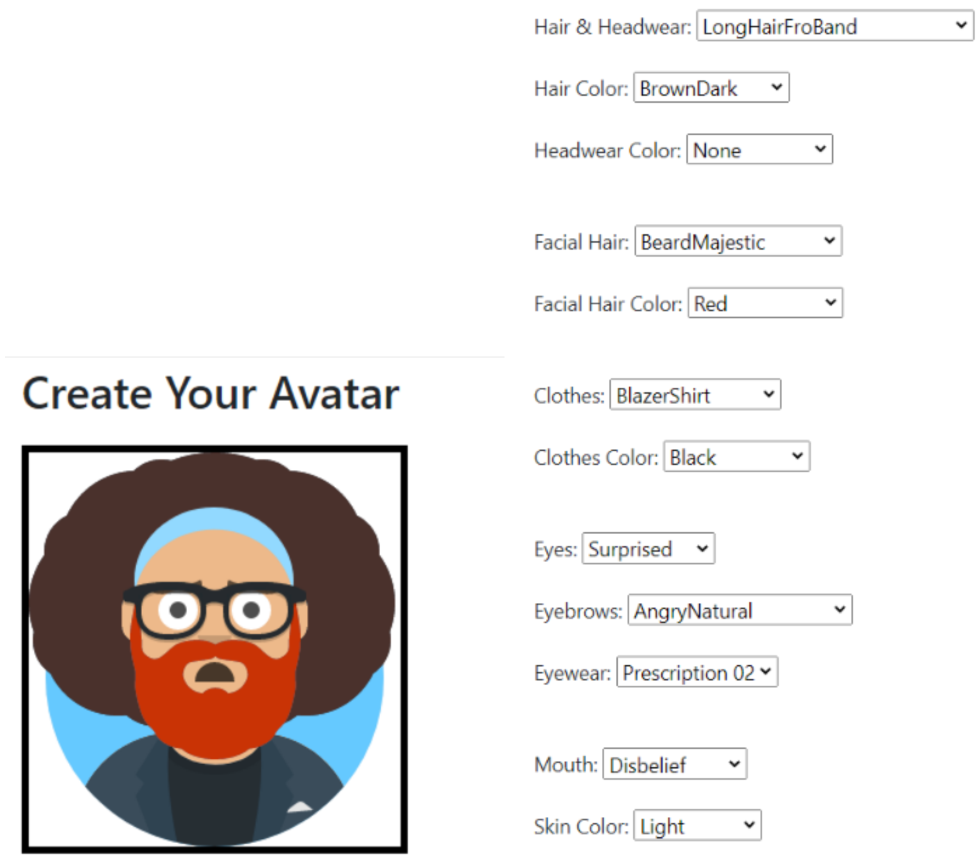


Figure 1.1: Experimental Interface, Avatar Customization. The image on the right shows the different attribute categories and dropdown menus available to subjects. The image on the left shows the avatar that corresponds to the specific attributes chosen in the image on the right.

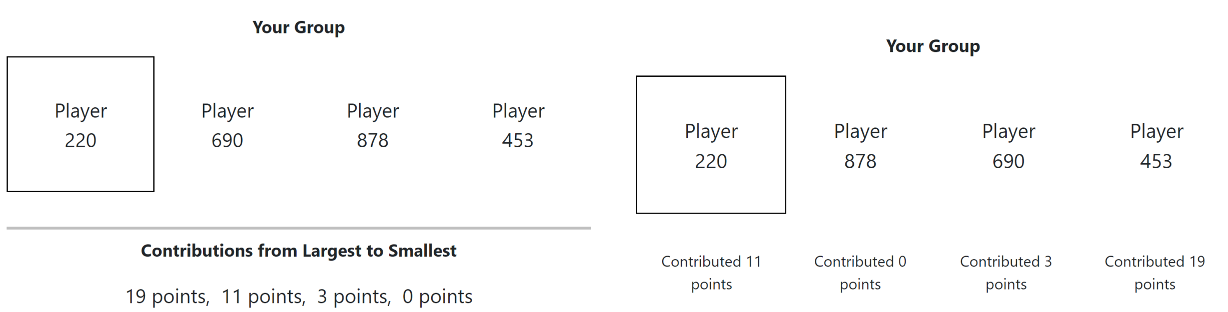


Figure 1.2: Experimental Interface, Round Results. *Partial Attribution* (left) and *Full Attribution* (right).

	Aggregate Only	*Partial Attribution*	*Full Attribution*
No Identifier	10		
Number		10	12
Avatar		11	14
Photo		10	10

Table 1.1: Number of Groups Assigned to Each Condition.

The combinations of different Identifiers and different Attribution schemes create different levels of anonymity across treatment conditions. A subject who is identified by their photograph in the game can be definitively identified by others outside of the lab, whereas an avatar typically will (and always can) look different from the corresponding subject, and a number never allows a physical identification. Similarly, one can definitively attribute a contribution decision to an identifier under *Full Attribution*, whereas one can only do so probabilistically under *Partial Attribution*. In the experiment, the condition with the most anonymity is the *Baseline* condition: subjects have no identifiers whatsoever, and individual contribution choices are not revealed. The least anonymous condition is *PhotoFull*: actions taken inside the lab can (1) be attributed with certainty to an identifier that (2) allows the subject to be identified outside the lab. In all remaining conditions, one of these links is weakened, so the level of anonymity sits somewhere in between *Baseline* and *PhotoFull*. For example, the *NumberFull* condition still attributes actions taken in the lab to an identifier with certainty, but the identifier itself—the random number—doesn't serve to identify the subject beyond the experiment. On the other hand, the *PhotoPartial* condition identifies subjects in a way that extends outside the lab, but does not allow actions taken in the lab to be attributed with certainty to the identifier.

One treatment condition was assigned to each session, and subjects within a session were randomly assigned to groups of four. Results from a power analysis dictated that each condition have at least 10 groups. Some conditions had more than 10 groups due to higher-than-expected turnout by subjects who signed up for some experimental sessions. Table 1.1 summarizes the number of groups assigned to each treatment.

The number of groups in each session varied based on the number of subjects that signed up and turned out. For *Baseline*, *Number*, and *Avatar* sessions, between two and five groups participated in each session: this ensured that subjects could not identify their group members outside the lab with certainty, just by seeing who else was in the lab with them. Between one and three groups participated in each *Photo* session. Recall that each subject belongs to and interacts with only one group for the duration of the experiment—groups are never reshuffled or remade—so the total number of subjects in a session does not affect any strategic aspect of the game. In total, thirty sessions were run.

The experiment was conducted in person at the Columbia Experimental Laboratory for Social Sciences (CELSS) between September 2022 and March 2023. 308 subjects were recruited from the CELSS subject pool using the Laboratory's ORSEE recruitment system ([46]). Most subjects were undergraduate students at Columbia University or Barnard College. The experiment was programmed in oTree ([47]) and subjects accessed the experimental interface on desktop computers at CELSS. Each experimental session lasted about 60 minutes. Average subject earnings were $22.07, including a $5 show-up payment. Instructions were read aloud to all subjects by the experimenter at the front of the room before the start of the game, and all subject questions were answered publicly.[31]

1.3 Theoretical Framework for Anonymity

A key finding of [3] is that attributing individuals' contribution choices to their photographs induces higher contributions in a finitely repeated public goods game. This finding is commonly understood to reflect the fear of being identified as a low contributor in the group; indeed, [13] find that revealing the two lowest contributors and their contributions between rounds raises contributions as much as revealing all players and their contributions. However, since the game is finitely repeated, adding identifiers or the individual contribution choices does not change the equilibrium prediction of free riding in each round. Moreover, even if players mistakenly believed that the game were infinitely repeated, fear of punishment in the game cannot explain the higher contributions,

[31] The Appendix contains representative instructions, as well as additional screenshots of the experiment interface.

since punishment cannot be individually targeted (unlike in a prisoners' dilemma).

It follows that we must consider other motivations that can explain this effect. In this section, I consider two channels through which anonymity (or a lack thereof) could affect contribution behavior: image concerns and punishment outside the game. I derive predictions for how these channels would influence contributions in the public goods game. While these channels are standard in the literature, they will not be sufficient to fully explain the results I present in Section 1.4.

1.3.1 Image Concerns

One possible consequence of introducing identifiers or revealing individual contribution choices is that subjects have image concerns within the game. Further, suppose subjects derive utility not only from their payoff in the public goods game, but also from their self image and social image within that interaction. I model self image utility as coming a warm glow effect of contributing à la [10] and [11], and I model social image utility as coming from the various social effects of contributing—recognition, prestige, shame. I introduce additional parameters capture the intensity with which players 'care about' their identifier and the probability with which their action is attributed to the identifier. In the model, these parameters capture the level of anonymity in the game, and serve to amplify or attenuate image utility. I assume all agents are symmetric, and write agent i's utility function as follows:

$$U_i(g_i, g_{-i}) = \pi_i(g_i, g_{-i}) + (1 + b)W(g_i) + bpR(g_i) \tag{1.2}$$

Let $g_i \in [0, Y]$ denote the amount that agent i contributes to the public good.[32] The first term π_i is agent i's payoff from the public goods game, as defined in Equation 1.1. The second and third terms capture the agent's self image utility and social image utility, respectively. $W : [0, Y] \to \mathbb{R}$ maps from contribution choice to self-image (warm glow) utility; analogously, $R : [0, Y] \to \mathbb{R}$ maps from contribution choice to corresponding social image (recognition) utility. I assume that

[32] I allow non-integer contributions here for tractability.

W and R are weakly increasing in g_i, concave, and smooth: the more the agent contributes, the stronger the warm glow, and the more recognition they receive.

The parameters b and p model the level of anonymity in the game. $p \in [0, 1]$ parameterizes the probability with which the agent's action can be attributed to their identifier. The parameter $b \geq 0$ captures the intensity with which the agent cares about, invests in, or identifies with their identifier. I normalize $b = 0$ when players do not have identifiers. In the absence of identifiers, $p = 0$. Note that b appears in both image utility terms: self image utility and social image utility are amplified by the extent to which the agent cares about their identifier. Moreover, note that b and p are complements in the social image utility term. Even if one is identified by their photograph (large b), there is little social image to be gained from contributing if the contribution is unlikely to be attributed to one's photograph (small p). Similarly, there is little social image to be gained from contributing if the contribution will be attributed to an identifier one cares little about (small b, large p).

In the experiment, $b = p = 0$ in the *Baseline* condition. Under *Full Attribution*, $p = 1$, and under *Partial Attribution*, $p = \frac{1}{3}$.[33] While impossible to ascribe values of b to a random number, cartoon avatar, or photograph, it seems reasonable to suppose that the following holds: $b_{\text{Number}} \leq b_{\text{Avatar}} \leq b_{\text{Photo}}$, with at least one inequality being strict.

As I show formally in the Appendix, sufficiently strong image concerns can induce positive contributions in equilibrium.[34] I state here two key results from the image concerns framework, and their associated predictions for the experiment.

Result 1. Equilibrium contributions g^* are increasing in b and in p: $\frac{\partial g^*}{\partial b} > 0$ and $\frac{\partial g^*}{\partial p} > 0$.

Prediction 1a. Contributions in the Baseline will be lower than contributions in all six treatment conditions.

Prediction 1b. $g^*_{\text{Baseline}} < g^*_{\text{Partial Attribution}} < g^*_{\text{Full Attribution}}$, for all identifiers.

[33] I state these probabilities from the perspective of a subject in the game. Assuming that subjects recall their own contribution and identifier, there remain three contribution choices that need to be attributed to identifiers. Under *Partial Attribution*, the probability that a contribution choice will be attributed to the correct identifier is $\frac{1}{3}$.

[34] I show formally in the Appendix that equilibrium contributions can be strictly positive, and can be as large as the endowment Y under certain conditions.

Prediction 1c. $g^*_{\text{Number}} \leq g^*_{\text{Avatar}} \leq g^*_{\text{Photo}}$, with at least one inequality being strict, for all Attribution schemes.

Result 2. g^* has increasing differences in (b, p), since b and p are complements.

Prediction 2a. $(g^*_{\text{PhotoFull}} - g^*_{\text{PhotoPartial}}) > (g^*_{\text{NumberFull}} - g^*_{\text{NumberPartial}})$. Contributions are more sensitive to a change in p when b is larger.

Prediction 2b. $(g^*_{\text{PhotoFull}} - g^*_{\text{NumberFull}}) > (g^*_{\text{PhotoPartial}} - g^*_{\text{NumberPartial}})$. Contributions are more sensitive to a change in b when p is larger.

1.3.2 Punishment outside the Game

Another possible consequence of introducing identifiers or revealing individual contribution choices is that it enables players to be punished outside of the public goods game itself. Imagine a group of friends which gathers one evening to play a board game. Actions taken in the game may spill over into their larger friendship: a person who behaves selfishly in the game may be shunned by the other friends afterward, even after the game itself is over. I model this intuition in my setting as follows: I consider an infinite horizon supergame, in which a finitely repeated public goods game is followed by an infinitely repeated, costly punishment game. In every period of this punishment game, each player can punish as many (or few) of their group members as they want to; punishing another player, as well as being punished, are both costly. Can the threat of being punished deter free riding in the preceding public goods game? I show formally in the Appendix that the answer is yes.

This framework is relevant to the experiment because the equilibrium of the supergame depends on whether the identity of a free rider in the public goods game can be common knowledge. If a free rider's identity can be commonly known, the other players are able to select the equilibrium that inflicts maximal punishment on the free rider. If not, then the free rider escapes maximal punishment with strictly positive probability. Thus, when a free rider's identity cannot be commonly known, the expected punishment from free riding is strictly lower. This further implies that there exists some discount factor that can sustain positive contributions when a free rider's identity can be commonly known, but not when a free rider's identity cannot be known; moreover, the reverse does not hold.

When can the identity of a free rider, or deviator, be commonly known in the public goods game? Note that for any given profile of contribution strategies in the public goods game, observing the aggregate contribution after each round is sufficient to reveal the *presence* of a deviator.[35] But only when a deviation can (1) be attributed with certainty to an identifier that (2) allows the deviator to be identified outside the game, can the deviator's identity be common knowledge. In the experiment, only in the *PhotoFull* treatment are conditions (1) and (2) both satisfied.

Result 3. There exists some discount factor $\delta \in (0,1)$ that can sustain positive contributions in the public goods game if (1) a deviation can be attributed with certainty to an identifier that (2) allows the deviator to be identified outside the game, but cannot sustain positive contributions otherwise. The reverse does not hold.

Prediction 3. $g^*_{\text{PhotoFull}} > g^*_{\text{All Other Conditions}}$.

1.4 Experimental Results

1.4.1 Overview

Average contributions aggregated across all rounds are displayed in Figure 1.3, while Figure 1.4 shows the average contribution over time. While subjects do not play the Nash equilibrium of contributing zero, contributions decline over time in all treatment conditions. Both observations are consistent with behavior of subjects in past experiments with public goods games.

1.4.2 The Treatment Effect of Identifiers and Attribution

Does introducing identifiers or revealing individual contribution choices increase contributions? Figure 1.3 shows that the answer is yes. Average contributions are higher in all six treatment conditions than in the *Baseline* condition. Moreover, Figure 1.4 shows that contributions under

[35] For example, consider a profile of strategies that dictates that all N group members contribute their entire endowment. The presence of a deviator is revealed if the aggregate contribution is observed to be less than NY.

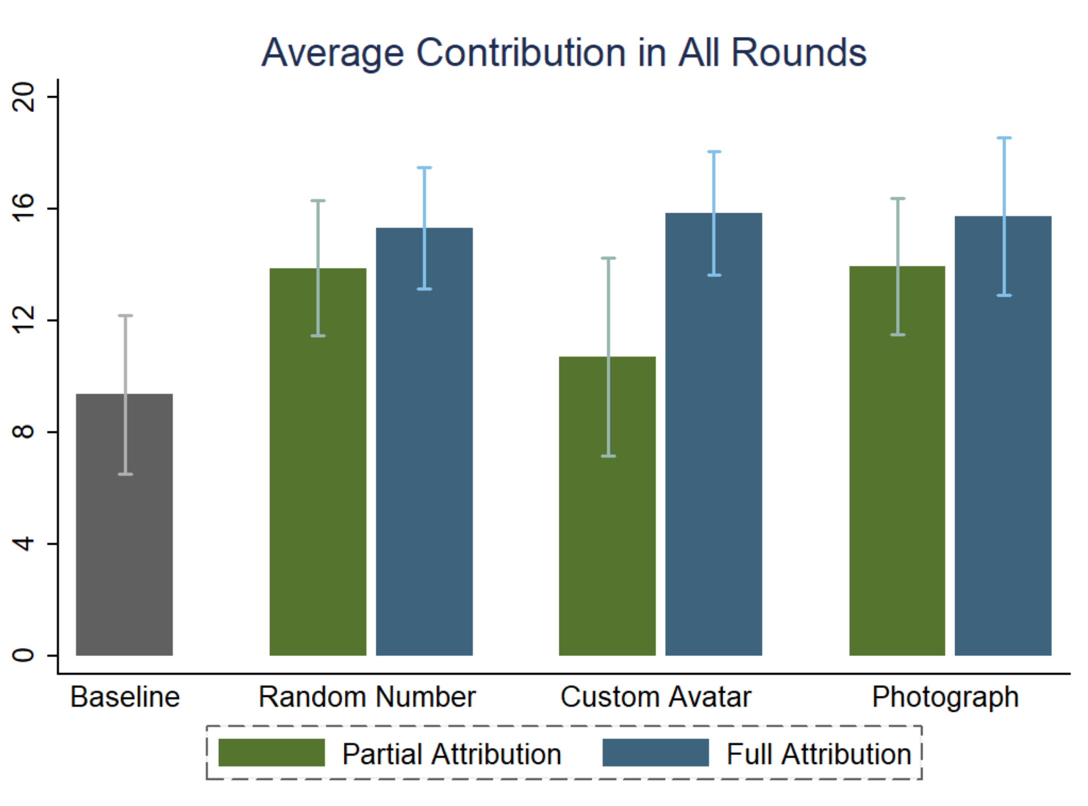


Figure 1.3: Average Contribution by Treatment. Error bars show a 95% confidence interval, with standard errors clustered at the group level.

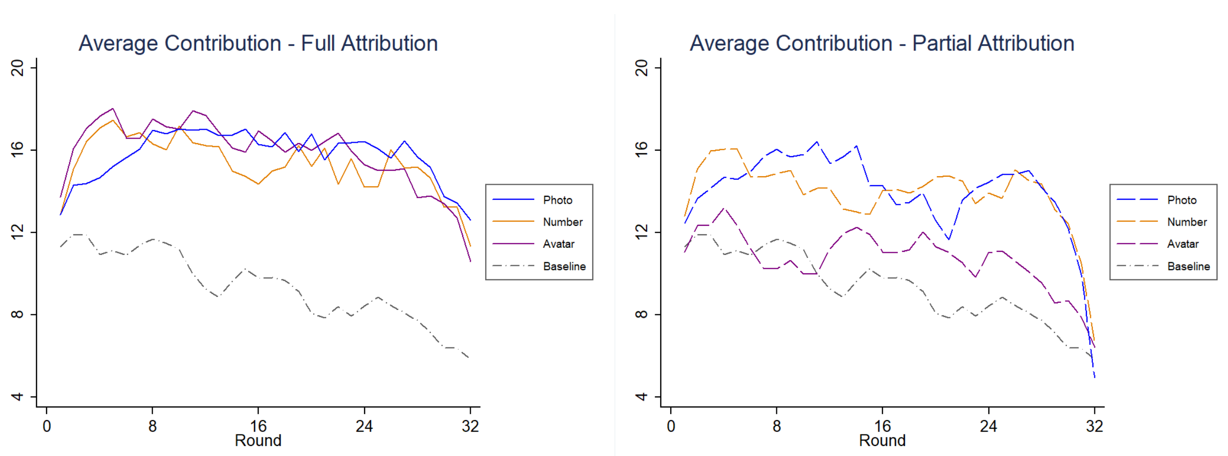


Figure 1.4: Average Contribution over Time.

Condition	Average Contribution (as % of Endowment)	Mann Whitney Test Statistic (vs Baseline)	p-value
Baseline	47%	-	-
NumberPartial	69%	2.26	0.023
NumberFull	77%	2.57	0.010
AvatarPartial	54%	0.63	0.526
AvatarFull	79%	2.86	0.004
PhotoPartial	70%	2.04	0.041
PhotoFull	79%	2.79	0.005

Table 1.2: Average Contribution and Mann Whitney Test by Treatment.

Full Attribution lie strictly above contributions in the *Baseline* condition, even in the final round of the game. To test whether contributions in the treatment conditions are statistically significantly different from contributions in the Baseline, I perform a Mann Whitney U test, where each group of four subjects is treated as one observation.[36] The U test shows that the difference in average contributions is significant at the 5% level for five of the six treatment conditions (*AvatarPartial* being the exception). Table 1.2 summarizes the results.

As contributions in the treatment conditions are higher than in the *Baseline*, *Prediction 1a* from the image concerns framework is satisfied. On the other hand, contributions do not exhibit increasing differences. In fact, across all Identifiers, contributions stay equally high under *Full Attribution*. Hence, the data are not consistent with *Prediction 2a* and *Prediction 2b* from the image concerns framework. Moreover, contributions in *PhotoFull* are not significantly greater than all other conditions: *Prediction 3* from the punishment framework is also not satisfied.

To isolate the effect of changing Identifier (Attribution) on contributions, I pool the data from the treatment conditions across Attribution (Identifier). The data show that changing the level of Attribution has a stronger effect on contributions than changing Identifier. The left panel in Figure 1.5 shows average contributions for the different levels of Attribution, pooling across Identifiers; the right panel shows average contributions for the different Identifiers, pooling across levels of

[36]Recall that the public goods game is played in fixed groups: each subject had strategic interactions only with the other three members of their group. Thus, I treat each group as an independent observation.

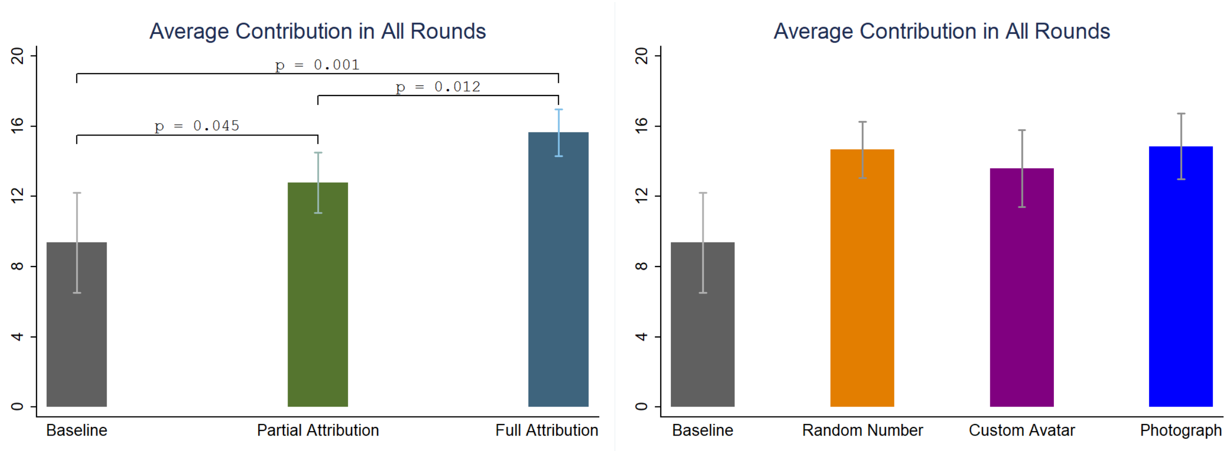


Figure 1.5: Average Contribution by Attribution (left) and by Identifier (right). Error bars show a 95% confidence interval, with standard errors clustered at the group level. p-values are shown for a Mann Whitney U test, where each group of four subjects is treated as one observation.

Attribution. Pooling across Identifiers, the average contributions in Baseline, Partial Attribution, and Full Attribution are all significantly different from each other at a 5% level: *Prediction 1b* from the image concerns framework is satisfied. On the other hand, when pooling across Attribution levels, average contributions in Number, Avatar, and Photograph are not significantly different from each other: the data are not consistent with *Prediction 1c*.

To understand what is causing these differences, I examine extreme behavior. How often do subjects contribute nothing or contribute their full endowment, and how does that vary across conditions? Table 1.3 reports the percentage of subject-round observations in which the contribution was zero; Table 1.4 reports the percentage of subject-round observations in which the entire endowment was contributed. There is no significant difference in the rates of zero contribution or full contribution between the three Identifiers, aggregating over Attribution schemes.[37] On the other hand, there are significant differences between the different Attribution schemes, aggregating over Identifiers. Most striking is the difference in rate of full contribution: in the *Baseline*, subjects contributed their full endowment in 16.6% of rounds. That figure rises to 38.6% in *Partial Attribution* rounds ($p = 0.048$ when compared to *Baseline*), and further to 62.3% in *Full Attribution* rounds

[37]I conduct a Mann Whitney U test to compare the rate of zero contribution or full contribution, where each group is treated as one observation. For each group, I calculate the fraction of the 128 subject-round observations (4 subjects per group play 32 rounds each) that have zero contribution or full contribution.

	Aggregate Only	*Partial Attribution*	*Full Attribution*	Total
No Identifier	14.3%			14.3%
Number		7.2%	11.7%	9.6%
Avatar		14.6%	7.1%	10.4%
Photo		11.1%	3.8%	7.5%
Total	14.3%	11.1%	7.7%	9.9%

Table 1.3: Percentage of Subject-Round Observations with Zero Contribution.

	Aggregate Only	*Partial Attribution*	*Full Attribution*	Total
No Identifier	16.6%			16.6%
Number		45.1%	60.9%	53.7%
Avatar		30.1%	66.4%	50.4%
Photo		41.5%	58.1%	49.8%
Total	16.6%	38.6%	62.3%	46.8%

Table 1.4: Percentage of Subject-Round Observations with Full Contribution.

(p = 0.005 when compared to *Partial Attribution*, $p < 0.001$ when compared to *Baseline*).[38] This suggests that the increase in average contributions in the treatment conditions is primarily driven by increased rates of full contribution when individual contribution choices are made common knowledge.

Coming back to the theoretical framework presented in Section 1.3, the experimental data are not consistent with the punishment framework. Contributions in the *PhotoFull* condition are not significantly higher than in all other conditions; in fact, contributions in *RandomFull* and *Avatar-Full* are just as high, even though subjects cannot be identified outside the lab with a number or avatar. The data are also not fully consistent with the image concerns framework either. While contributions are increasing in p as predicted, contributions are not increasing in b, nor do they exhibit increasing differences in (b, p).

While one can interpret the data to say that image concerns are not at play, an alternative interpretation is that b_{Number}, b_{Avatar}, and b_{Photo} are very close, if not equal, to each other—but still greater than zero. Consider a modification of the image concerns framework where, instead

[38]The differences across Identifiers in rate of zero contribution were not as stark or statistically significant. The decrease in zero contribution rate from *Baseline* to *Partial Attribution* was not significant (p = 0.267). The decrease in zero contribution rate from *Partial Attribution* to *Full Attribution* was significant at the 10% level (p = 0.064). The decrease in zero contribution rate from *Baseline* to *Full Attribution* was significant at the 5% level (p = 0.022).

of assuming $b_{\text{Number}} < b_{\text{Avatar}} < b_{\text{Photo}}$, the assumption was that there is some $b > 0$ as long as any identifier exists. While such an assumption is perhaps unintuitive, it would allow the image concerns framework to fit the data better. Taken in conjunction with the [3] result, these results suggest that the role of an identifier in inducing higher contributions is not to identify the physical person, but rather to provide something that an action can be attributed to, whether with certainty or probabilistically.

1.4.3 Dynamics of Contributions

Given the results so far, what could be driving contributions? In this section, I focus my analysis on the difference in contributions across Attribution schemes. Note that when subjects are provided individual contribution choices, they are also given more information than if they only had the aggregate contribution. For example, when a subject sees the contribution choices of their group members, they know where their contribution ranks within the group; in most cases, they could not know this with only the aggregate contribution.[39] The analysis in this section asks: *Are subjects reacting to the information provided by the individual contribution choices under Partial and Full Attribution?* In particular, I consider a purely reactive model: I assume that subjects are backward-looking and respond simply to behavior in the previous round, without anticipating future reactions.

To address this question, I analyze the dynamics of contributions over the course of the experiment: I seek to identify the factors that explain how a subject's contribution varies from one period to the next, and whether these factors differ across Attribution schemes. I focus on the data aggregated at the Attribution level (*i.e.*, pooled across Identifiers). For each level of Attribution, I separately regress player i's contribution in round t on regressors from the previous round, $t - 1$. The first four regressors correspond to pieces of aggregate information. *OwnContrib* is the subject's private contribution, and *SumOtherContribs* is the sum of the contributions made by the other members of the subject's group. *AboveMean* (*BelowMean*) is an indicator variable for whether the

[39] There are a few edge case exceptions. For example, if the aggregate contribution is zero, then it can be known that every player chose to contribute zero.

subject's contribution was above (below) the mean contribution by their group members.[40] The final three regressors, on the other hand, correspond to pieces of relative information that can only be known when the individual contribution choices are also revealed. *RankInGroup* is a variable that equals the number of group members who contributed strictly more than the subject. *SomeoneElseFreeRide* (*SomeoneElseFullContrib*) is an indicator variable for whether another member of the group contributed zero (contributed their full endowment).

I estimate a dynamic panel model using the generalized method of moments, specifically the [21] estimator. Table 1.5 displays the results.

As expected, one's contribution in the previous period (*OwnContrib.lag1*) has a significant, positive impact on one's contribution in the current period, and there is a significant, negative time trend in all conditions. Additionally, I find a significant, positive coefficient *RankInGroup.lag1* and a significant, negative coefficient on *SomeoneElseFreeRide.lag1* for the Partial and Full Attribution subjects. Simultaneously, this reduces the size of the positive coefficient on *BelowMean.lag1* for Partial and Full Attribution subjects. This suggests that subjects are, in fact, conditioning their future contributions not only on information they glean from the aggregate contribution, but also from the relative information they get from the individual contribution choices. Subjects significantly reduce their contribution in the subsequent period when at least one of their group members contributed nothing; yet, they significantly increase their contributions in proportion to the number of group members that contributed strictly more than them. Moreover, note that the coefficient on the final three regressors is not significant for the Baseline data. This serves as a placebo test: Baseline subjects cannot know where their contribution ranks, or whether someone else in the group free rode or contributed the full endowment. Indeed, the data suggest that these regressors do not predict contributions by Baseline subjects.

The patterns revealed in this analysis are broadly consistent with a number of models that have been used to explain contribution behavior in public goods games. [22] finds that subjects in a finitely repeated public goods game tend to "match" the median contribution from the previous

[40] Whether the mean includes or excludes the subject's own contribution is irrelevant.

	(1) Baseline	(2) Partial	(3) Full	(4) Baseline	(5) Partial	(6) Full	(7) Baseline	(8) Partial	(9) Full
OwnContrib.lag1	0.45***	0.60***	0.60***	0.50***	0.71***	0.72***	0.48***	0.73***	0.80***
	(0.11)	(0.05)	(0.06)	(0.12)	(0.06)	(0.07)	(0.12)	(0.06)	(0.08)
SumOtherContribs.lag1	0.03	0.08***	0.01	0.00	0.04	-0.03	0.03	0.01	-0.07**
	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)
AboveMean.lag1	-1.48*	-1.93***	-1.94***	-1.55*	-2.26***	-2.46***	-1.36	-2.19***	-2.46***
	(0.75)	(0.43)	(0.45)	(0.78)	(0.43)	(0.49)	(0.70)	(0.43)	(0.48)
BelowMean.lag1	1.43*	4.72***	5.68***	0.68	2.10***	1.75**	0.93	1.82***	1.57*
	(0.65)	(0.58)	(0.64)	(0.66)	(0.56)	(0.62)	(0.60)	(0.53)	(0.64)
Round	-0.09**	-0.08***	-0.07***	-0.10***	-0.07***	-0.06***	-0.10**	-0.07***	-0.05***
	(0.03)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.03)	(0.02)	(0.01)
RankInGroup.lag1				0.73	1.82***	2.35***	0.70	1.94***	2.31***
				(0.40)	(0.28)	(0.29)	(0.40)	(0.28)	(0.31)
SomeoneElseFreeRide.lag1							0.30	-1.74***	-2.56***
							(0.45)	(0.48)	(0.61)
SomeoneElseFullContrib.lag1							-0.93	-0.13	1.04
							(0.56)	(0.46)	(0.57)
Constant	5.82***	2.45*	5.94***	5.60***	1.87	5.47***	5.03**	2.83*	6.25***
	(1.50)	(1.05)	(1.27)	(1.55)	(1.12)	(1.35)	(1.67)	(1.11)	(1.36)
Observations	1200	3720	4320	1200	3720	4320	1200	3720	4320
Number of Subjects	40	124	144	40	124	144	40	124	144

Robust standard errors in parentheses. Estimation performed using the *xtabond* command in STATA.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.5: Arellano-Bond Estimation.

round, and that such behavior is most consistent with a model of reciprocity. In my experiment, I find that subjects who contribute more than the mean tend to reduce their contribution in the next period, while subjects who contribute less than the mean tend to increase their contribution. This behavior could also be interpreted in two other ways. First, this behavior could be taken as evidence of backward-looking inequality aversion: if subject derive disutility from contributing more (less) than the average group member, they would adjust their subsequent contribution down (up). Such behavior is consistent *prima facie* with the model in [9], though one key reversal in my data is that subjects who contribute less (earn more) than the mean respond much more strongly than subjects who contribute more (earn less) than the mean. The second interpretation of the contribution dynamics is that subjects use the feedback about individual contribution choices in order to learn and establish a norm for the group: this is particularly noticeable under *Partial Attribution* and *Full Attribution*, when subjects who contribute below the mean increase their contribution by a greater amount than by which subjects who contribute above the mean reduce their contribution.

Interestingly, the aforementioned effects are stronger under under *Full Attribution* than under *Partial Attribution*. Image concerns are one explanation of this: recall that, while the data are inconsistent with the image concerns framework for different Identifiers, the data are consistent with the image concerns framework regarding different Attribution schemes. One reading of the data, considering norm-setting and image concerns together, says that the effect observed under *Partial Attribution* captures the extent to which subjects use relative information to learn and establish a norm for the group. The additional effect observed moving from *Partial Attribution* to *Full Attribution* captures the image concerns associated with being seen as a relatively low contributor (*RankInGroup.lag1*) or as a "sucker" who contributed when someone else free rode (*SomeoneElseFreeRide.lag1*).

1.4.4 Avatars

In this section, I analyze two measures of subjects' connectedness to their avatars, and whether those measures are predictive of contribution behavior in the first round of the public goods game.[41] Additionally, I analyze whether subjects create avatars that resemble their physical persons. The first measure of connectedness is the amount of time the subject spent on customizing their avatar, and the second measure is the subject's self-reported intensity of feeling represented by their avatar. Figure 1.6 shows the avatars created by the 100 subjects randomly assigned to the Avatars condition.

While all subjects began customizing their avatars at the same time, each subject could decide for themselves when they were finished customizing their avatar by clicking a button at the bottom of the avatar customization screen; only when all subjects finished customizing their avatars did the public goods game begin. I calculate the amount of time spent by each subject on customizing their avatar as the difference (in seconds) between the time they clicked their button and the time they entered the avatar customization screen. Across the 100 subjects that were assigned to the Avatars condition, the mean amount of time spent on customizing an avatar was just over 3 minutes (189 seconds), with a minimum of 53 seconds and a maximum of 353 seconds.[42] The was no significant difference in customization time across Attribution: under Partial Attribution, average customization time was 188 seconds; under Full Attribution, average customization time was 190 seconds. However, the time spent on avatar customization does not have a significant effect on contributions: regressing own contribution in the first round on customization time and a Full Attribution dummy variable (and thus controlling for the Attribution scheme) yields a coefficient on customization time that is not significantly different from zero ($\beta = 0.005$, $p = 0.507$).

After the conclusion of the public goods game (but before being told their final earnings),

[41] Because contribution decisions in later rounds can be influenced by the contribution decisions in preceding rounds, I restrict the analysis to the first round only.

[42] While subjects were instructed that they had five minutes to finish customizing their avatars, they were allowed a bit of flexibility: subjects were given a "2 minutes remaining" announcement, followed by a "1 minute remaining" announcement, followed by an announcement asking subjects still customizing their avatar to wrap up. This ensured that each subject entered the game with an avatar that they were satisfied with.

Figure 1.6: Avatars created by subjects.

subjects self-report their race, gender, and ethnicity. Additionally, they rate their agreement with the following statement: *I felt like my avatar represented me.*[43] The rating was done on a 7 point Likert scale, where 1 was *Strongly Disagree*, 4 was *Neutral*, and 7 was *Strongly Agree*.[44] On average, subjects in the *Partial Attribution* condition agreed more with the statement (average rating = 3.88) than the subjects in the *Full Attribution* condition (average rating = 3.38), but this difference is not significant (p = 0.221 from a Mann Whitney U test). There is also no correlation between the amount of time a subject spent on their avatar and their representation rating.

Does the reported level of representation correlate with first round contributions? The evidence is, at best, very weak. Under *Partial Attribution*, I can reject the null hypothesis of independence at the 10% level.[45] I cannot reject independence under *Full Attribution*.

Do subjects create avatars that resemble their physical persons? I analyze the similarity of avatar to physical person in two ways. First, I compare the subject's self-reported race and ethnicity to the skin tone of their self-designed avatar.[46] Second, I manually code each avatar with a gender (male, female, neutral) and compare the avatar's gender to the self-reported gender of the subject who created that avatar. On both dimensions, there is evidence that some subjects create avatars that don't resemble their physical person. Only one subject chose the *Black* skin tone for their avatar, yet that subject self-identified as non-Hispanic White. There is a similar disparity between genders of avatars and the subjects who made them. Of seventeen subjects who self-identified as male, five designed avatars that were coded as female. Of 64 subjects who self-identified as female, six designed avatars that were coded as male and thirteen designed avatars that were coded as neutral gender. At the intersection of race and gender, five subjects self-identified as Black women. Only three of those five subjects created avatars that coded as female, and none of those

[43]Due to a server problem, the demographic data and agreement with representation statement was not collected for one session. The treatment in that session was *AvatarFull* and 16 subjects participated in that session.

[44]Subjects saw only the labels. They were not shown the numbers 1 through 7.

[45]The Spearman rank correlation coefficient between reported level of representation and level of contribution in the first round for Partial Attribution subjects is 0.28; the null hypothesis of Independence is rejected at the 10% level (p = 0.064). Because the reported level of representation is a categorical variable, I use the Spearman correlation coefficient instead of the Pearson correlation coefficient.

[46]The available skin tone options, from lightest to darkest, were: *Pale*, *Light*, *Yellow*, *Tanned*, *Brown*, *DarkBrown*, *Black*.

Figure 1.7: Avatars created by subjects who self-identified as Black women.

subjects chose the *Black* skin tone for their avatars: Figure 1.7 presents the five avatars created by these five subjects.

1.5 Conclusions

As our interactions shift increasingly to the digital world, it has become important to understand how its peculiarities affect our behavior. The digital world allows individuals to identify themselves with a wider range of identifiers than is available to individuals in the physical world. Moreover, the digital world allows interactions among individuals without identifiers at all, which further complicates how digital actions are attributed to the physical persons behind the screen.

The experimental data confirm that identifying individuals and providing information about individual actions work in conjunction to increase contributions in a public goods game. Surprisingly, however, the positive effect persists when actions are attributed to identifiers that are unable to identify the physical person, when actions can only be probabilistically attributed to the physical person, and even when actions can only be probabilistically attributed to identifiers that are unable to identify the physical person.

Pitting the two channels against each other, on the hand, the data show that variations in the Attribution scheme have a larger effect on contributions than variation in Identifier. This difference is driven by higher rates of subjects contributing their full endowment under *Full Attribution* than under *Partial Attribution* than under *Baseline*. Dynamic estimation of the contribution decisions suggests that subjects react to "relative" information when available: as subjects proceed through the game, they attempt to figure out, establish, and follow a norm. Indeed, many *Full*

Attribution groups successfully coordinate on contributing their entire endowment, even without a coordinating device or central authority.[47]

Interestingly, much of the existing policies targeting anonymity in digital settings have revolved around users having identifiers that don't pin down their physical persons. Many online communities, from Facebook[48] to Statalist (the official Stata forum)[49], require users to use their "real" names as their identifiers. Yet, these policies have faced substantial resistance: for instance, Facebook has gotten pushback about its "real name" policy from users of various ethnicities, from transgender users, from users who needed their identities protected.[50] [51] [52] However, the results from this experiment suggest that tech companies designing the digital world, and policymakers seeking to regulate it, may do better to emphasize a system where a user's history of actions is logged and attributed to some identifier, as opposed to a "real name" policy where users must use their real name.

The digital world is still young, and there remain many decisions to be made about what it will look like. Given the myriad anecdotes of troubling online behavior, inducing individuals to behave prosocially should be a priority in the design of digital spaces where many people gather and interact. Although much remains to be studied, the results of this paper suggests a path forward to ensuring that people do good in the digital world.

47 This is particularly encouraging for proponents of decentralized governance in digital settings.

48 "Facebook is a community where everyone uses the name they go by in everyday life." https://www.facebook.com/help/112146705538576.

49 "You are asked to post on Statalist using your full real name [...] . Giving full names is one of the ways in which we show respect for others and is a long tradition on Statalist. It is also much easier to get to know people when real names are used." https://www.statalist.org/forums/help.

50 https://www.telegraph.co.uk/news/newstopics/howaboutthat/2632170/Woman-called-Yoda-blocked-from-Facebook.html

51 https://www.cnn.com/2014/09/16/living/facebook-name-policy/index.html

52 https://www.huffpost.com/entry/pakistans-religious-extremism_b_9577338

Chapter 2: Avatars, Self Image, and Social Image

2.1 Introduction

One's self image and social image have long been considered important, economically meaningful influences on behavior and outcomes. Yet, it is unclear how one's image translates to digital interactions, where individuals need not identify and represent themselves with their "real" name, face, or body. The internet has enabled individuals to interact with each other in digital spaces, and advancements in augmented and virtual reality technologies has made these interactions increasingly lifelike and sophisticated. Whereas people were limited in the past to text-based usernames or rudimentary images, technology today allows individuals to customize sophisticated avatars to use in digital settings. In contrast to how individuals are identified and represented in physical interactions, avatars need not resemble the physical appearance of the individuals behind the screen. This distinction underlies the main question that this paper seeks to address: *To what extent does an avatar substitute for one's physical appearance in inducing image considerations?* As more interpersonal interactions take place online, it is imperative to understand how individuals use avatars as a form of self-representation and of representation to others.

To address this question, I design and run two separate (but related) experiments to measure the image considerations induced by avatars. I embed avatars into a one-shot dictator game, a simple interaction designed to reduce strategic considerations and isolate the effect of the avatars. Experiment S is designed to measure the degree to which seeing one's avatar heightens self image considerations. In Experiment S, Senders are given five minutes to customize a cartoon avatar before deciding how much of a $10 endowment to send to an anonymous Receiver. The avatars underpin the main experimental variation: in the spirit of the [14] experiment, treated Senders 'face' the avatar they created when deciding how much money to send. Control Senders, on the other hand, do not see their avatar again. Experiment R, on the other hand, is designed to examine the social image implications of the outward appearance of an avatar. In Experiment R, Receivers customize cartoon avatars that will be shown to the Sender. The main experimental variation comes in the timing: a control Receiver creates their avatar *before* learning that their avatar will be shown to a Sender in a dictator game, whereas a treated Receiver knows how their avatar will be used before they create it.

In Experiment S, I find that treated Senders send more money than control Senders on average, but that the difference is not statistically significant at the 5% level. On the other hand, I find in Experiment R that treated Receivers receive significantly more money than do control Receivers. To better understand the effect of the treatment on avatar appearance, I also ask subjects to rate avatars created by others on eight dimensions, each characterized by an opposing pair of words (e.g., *Attractive - Unattractive* or *Female - Male*). I find no significant impact of the treatment on avatar appearance in Experiment S, which is to be expected: the treatment effect in Experiment S does not interact with the avatar creation whatsoever. However, I do find an effect of the treatment on avatar appearance in Experiment R: treated Receivers created avatars that were rated as significantly more attractive, female, honest, and smart than control Receivers, even when controlling for demographic characteristics of the Receivers themselves. That said, the treatment effect on avatar appearance (measured by the ratings) doesn't explain the treatment effect on money received: variation in any given rating dimension does not explain variation in money received at a statistically significant level. Ultimately, this suggests that the treatment effect on avatars in Experiment R works holistically, as opposed to through a particular visual attribute or aspect of appearance.

2.1.1 Literature

Many theories incorporate the impact of image considerations into a decisionmaker's calculus.[1] Moreover, empirical studies have established the fact that individuals are treated differently by others depending on their physical appearance. For example, [50] show that beautiful people

[1]For example, [15], [48], and [49] incorporate image concerns into signaling models.

earn higher wages in the labor market than average-looking or plain-looking individuals. Lab experiments have also found similar results. [51], as well as [52], find a similar "beauty premium" in a dictator game played in a controlled laboratory setting. [53] also find evidence of the beauty premium in the context of a public goods game. Even simple social cues like a smile can affect behavior: [54] find that responders in an ultimatum game are more likely to accept the proposal if the proposer is smiling. Research has also found that the act of seeing oneself can affect one's decision in an interpersonal interaction. [14] reports the results of an experiment in which subjects decide whether or not to inflict a painful electric shock on an anonymous participant in return for money: he finds that subjects who see a video feed of themselves on the decision screen inflict the shock significantly less often than subjects who see no stimulus or a neutral stimulus on the same screen. [55] similarly find an effect of self image in the context of charitable giving. In their experiment, subjects decide how to allocate 10 pounds between themselves and either a charity or the experimenters. After the allocations are decided, subjects are informed which decision counts toward their payoff and are subsequently given the option to "opt out" of their decision and keep 10 pounds for themselves. They find that approximately one-third of subjects choose to opt out and ascribe the initial giving as motivated by factors other than pure altruism, such as self image concerns.

This paper contributes primarily to a literature that studies how individuals behave in settings where they are identified by avatars. There is evidence that individuals consider the types of people they will interact with as well as the nature of the interactions themselves when customizing their avatars in virtual worlds. [41] document this effect in an experiment on users on the Yahoo! Avatars site, and [42] find a similar effect using observational data, survey responses, and interviews of users in the *Second Life* virtual world. Recent work in economics finds a similar pattern. [43] find, in their experiment, that female subjects choose female avatars significantly less often when competing to be "hired" for a math task than when the nature of the task is unknown. [44] also find a gender effect in a marketplace experiment: buyers trust female avatars more than male avatars. In response, male sellers are more likely to select female avatars than vice versa.

The paper proceeds as follows. Section 2.2 describes the designs of Experiments S and R. Section 2.3 presents the results. Section 2.4 concludes. Experimental instructions appear in the Appendix.

2.2 Experimental Design

The study consists of two separate, but related experiments, denoted Experiment S and Experiment R. In both experiments, subjects face two main phases:

1. A one-shot Dictator Game, either as a Sender or a Receiver; and

2. An Avatar Task, either as a Creator or a Rater.

I begin by describing the two phases, followed by the Experiments.

2.2.1 The Dictator Game

The game is a one-shot dictator game in which two players divide an endowment of \$10. Subjects were randomly assigned to the role of Sender or Receiver, with equal probability.[2] The Sender (her) chooses an integer amount of dollars $s \in \{0, 1, ..., 10\}$ for her consumption, and the Receiver (he) receives the remaining $10 - s$ dollars for his consumption. The unique Nash Equilibrium in this game is for the Sender to keep the entire endowment for her consumption, and pass along nothing to the Receiver: $s^* = 10$.

2.2.2 The Avatar Task

Creators

Subjects assigned to Create were instructed as follows: "Use the dropdown menus to design the avatar that will represent you in the study." Each dropdown menu controlled a different visual attribute of a cartoon avatar, ranging from hairstyle to skin color. The avatar customization

[2]Neutral terms were used in the instructions for the experiment: the Sender was referred to as "Person A" and the Receiver as "Person B."

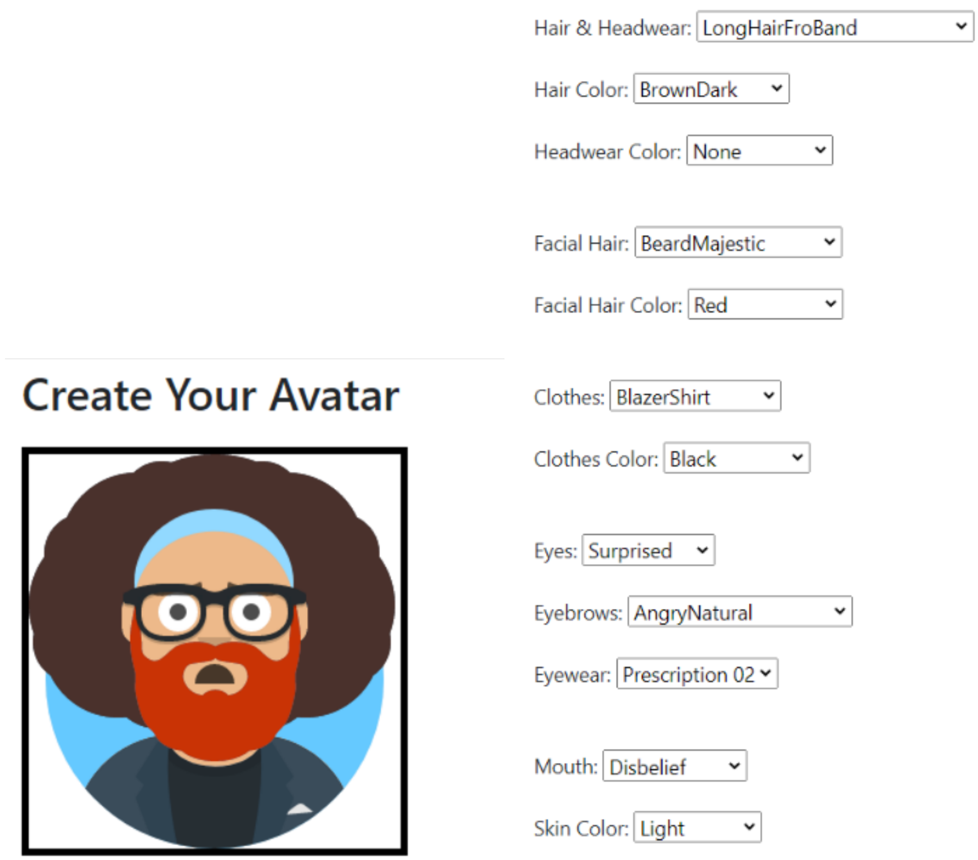


Figure 2.1: Interface for the Avatar Creation Task.

screen was coded in a way that allowed subjects to see, in real time, how changing a particular attribute would affect the appearance of the avatar. Hence, subjects could experiment with different attributes—and see how they would look—before finalizing their avatar. Subjects were forced to spend at least 150 seconds on the customization screen. Figure 2.1 shows screenshots of the avatar customization interface.

Raters

Subjects assigned to the Rating Task were asked to rate fifteen avatars on eight dimensions. Each dimension was characterized by an opposite word pair, such as *Attractive* - *Unattractive*, and a six-point scale ([56]). Subjects were instructed as follows: "For each word pair, please find the word that best fits the avatar. Please pick how well you think that word fits the avatar." Figure 2.2

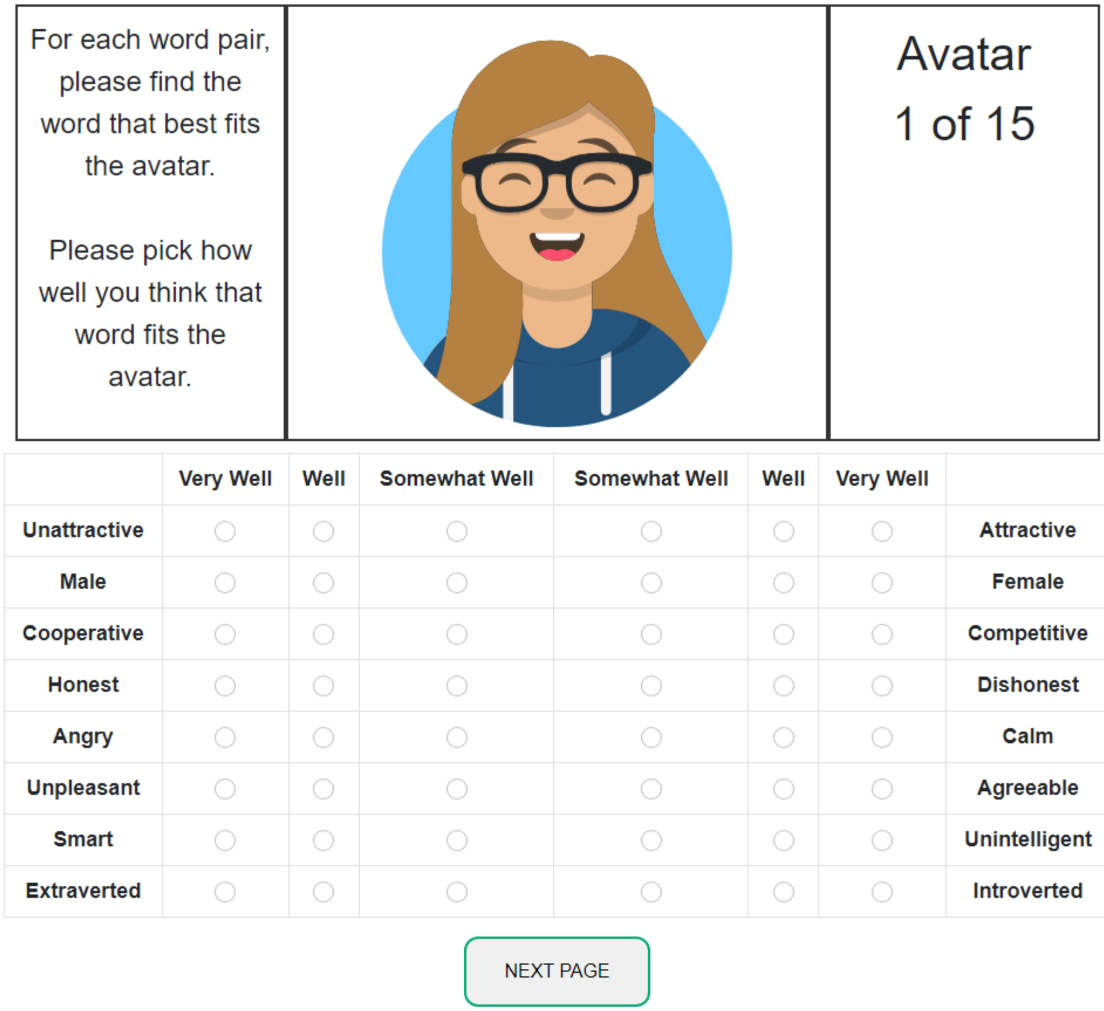


Figure 2.2: Interface for the Avatar Rating Task. The left-right order of each word pair is randomly determined in each round.

shows a screenshot of the avatar rating interface.

For each word pair, I code a rating of *Very Well* on the positive word (e.g., *Attractive*) as 6, and a rating of *Very Well* on its negative counterpart (*Unattractive*) as 1. Table 2.1 shows how the positive-negative distinction is coded for all word pairs.

2.2.3 Experiment S: Senders Create, and Receivers Rate

In Experiment S, subjects play the Avatar Task first, followed by the Dictator Game. Each subject is assigned to one of two joint roles:

“Positive” (*Very Well* = 6)	“Negative” (*Very Well* = 1)
Agreeable	Unpleasant
Attractive	Unattractive
Calm	Angry
Cooperative	Competitive
Extraverted	Introverted
Female	Male
Honest	Dishonest
Smart	Unintelligent

Table 2.1: Positive-Negative Coding for Eight Rating Dimensions.

1. Sender in the Dictator Game, and Creator in the Avatar Task; or

2. Receiver in the Dictator Game, and Rater in the Avatar Task

The main experimental manipulation occurs on the subjects assigned to the first role (Sender-Creator). In the treatment condition, subjects will see their avatar when deciding how much money to allocate to send in the dictator game. In the control condition, on the other hand, subjects do not see their avatar after they create it.

Subjects assigned to the second role (Receiver-Rater) are shown fifteen avatars, randomly chosen from the avatars created by subjects assigned to the Sender-Creator role. After rating the fifteen avatars, these subjects proceed to the dictator game and take no further action. By eliciting ratings for multiple avatars, I am able to collect multiple, independent ratings for any given avatar.

The main outcome of interest in Experiment S is whether subjects who see their own avatar send more money in the dictator game than subjects who don't see their own avatar. There are two secondary outcomes of interest in Experiment S. First, whether there are differences in ratings on any of the eight dimensions between avatars created by treated subjects and avatars created by control subjects. Second, whether there is a relationship between any of the eight dimensions and amount of money passed on in the dictator game.

2.2.4 Experiment R: Receivers Create, and Senders Rate

In Experiment R, roles are reversed. Each subject is assigned to one of two joint roles:

1. Receiver in the Dictator Game, and Creator in the Avatar Task; or

2. Sender in the Dictator Game, and Rater in the Avatar Task

The main experimental manipulation occurs in the timing of the game for subjects assigned to the first role (Receiver-Creator). In the treatment condition, subjects are given the instructions for the dictator game before creating their avatar, In particular, treated Receivers are told: "On the next page, you will design an avatar to represent yourself in this study. [The Sender] will be shown this avatar when they make the money allocation decision." In the control condition, the timing is reversed: subjects create their avatar before being told the context in which the avatar will be used.

Subjects assigned to the second role (Sender-Rater) first play three rounds of the dictator game. In each round, they are matched with a different Receiver-Creator, and are shown that Receiver's avatar when making their allocation decision.[3] Subsequently, each Sender-Rater is shown fifteen avatars, randomly chosen from the avatars created by subjects assigned to the Receiver-Creator role.[4]

There are two main outcomes of interest in Experiment R. First, whether treated Receiver-Creators receive more money from Senders than control Receiver-Creators. Second, whether avatars created by treated Receiver-Creators are more homogeneous on any of the eight dimensions than avatars created by control Receiver-Creators.

2.2.5 Implementation Details

Both experiments were coded using oTree ([47]). Both experiments were conducted online in February and March 2024, with 994 participants recruited on mTurk via CloudResearch[5] ([57, 58]). On average, Sender-Creators in Experiment S and Receiver-Creators in Experiment R spent about 9 minutes completing the experiment. Receiver-Raters in Experiment S and Sender-Raters

[3]Senders are told, before starting the dictator game, that "[w]hen you make your money allocation decision, you will see an avatar that [the Receiver] created to represent themself in the study."

[4]The three avatars that appeared in the preceding dictator game are excluded from random selection in the rating phase.

[5]Using CloudResearch, I restricted my sample to subjects residing in the United States who had previously completed at least 100 mTurk tasks (HITs) with at least a 90% approval rate.

	Experiment S	Experiment R
Sender	243 [Creator]	446 [Rater]
Receiver	59 [Rater]	246 [Creator]

Table 2.2: Number of Subjects Recruited, by Experiment and Role in Dictator Game.

in Experiment R spent about 12 minutes completing the experiment. Table 2.2 summarizes the number of subjects recruited for each role in each Experiment. Of the 243 Sender-Creators in Experiment S, 108 were assigned to the Control group and 135 to the Treatment group. Of the 246 Receiver-Creators in Experiment R, 118 were assigned to the Control group and 128 to the Treatment group.[6]

All subjects were paid $1 for completing the study, and were eligible to be selected for a bonus payment with 20% probability. If selected for bonus payment, the subject additionally earned their payoff in the Dictator Game, up to $10. In Experiment S, 59 Senders were randomly selected for bonus payment. In Experiment R, 51 Receivers were selected for bonus payment; Senders that matched with at least one of these Receivers were given a bonus payment.

In addition to the two main tasks of the experiment, all subjects completed a short, unincentivized demographic questionnaire.

2.3 Results

I first present balance tables of demographic characteristics for Creators in both experiments (Sender-Creators in Experiment S; Receiver-Creators in Experiment R), then discuss two main results, followed by results of exploratory analyses.

Table 2.3 (Table 2.4) reports average values or proportions of six demographic variables for Creators in Experiment S (Experiment R).[7]

[6] The computer program randomly assigned each Sender-Creator in Experiment S (Receiver-Creator in Experiment R) to Treatment or Control with equal probability.

[7] *isParent* is an indicator that takes value 1 for a subject who reports, for their parental status, that they are a parent. *isMale* is an indicator that takes value 1 for a subject who reports, for their gender, that they are a man. *isWhite* is an indicator that takes value 1 for a subject who reports, for their race, that they are white. *isLeft* (*isRight*) is an indicator that takes value 1 for a subject who reports, for their political views, that they are Left or Somewhat Left (Right or Somewhat Right).

Experiment S	Age (yrs)	isParent	isMale	isWhite	isLeft	isRight
Control (N = 108)	40.8	42.3%	57.5%	76.6%	40.0%	34.3%
Treatment (N = 135)	40.5	49.6%	64.1%	84.1%	46.0%	32.5%

Table 2.3: Demographic Characteristics of Creators in Experiment S.

Experiment R	Age (yrs)	isParent	isMale	isWhite	isLeft	isRight
Control (N = 118)	40.5	50.9%	67.5%	80.0%	48.7%	35.7%
Treatment (N = 128)	41.9	52.3%	57.8%	83.6%	47.6%	35.7%

Table 2.4: Demographic Characteristics of Creators in Experiment R.

2.3.1 Amount Sent (Received) by Creators in Experiment S (Experiment R)

The main outcome of interest in Experiment S was whether treated Senders, who saw their own avatars when making the money allocation decision, sent more money to Receivers than control Senders, who did not see their own avatars after designing them. Figure 2.3 shows that treated Senders do pass on more money: on average, treated Senders sent $3.42 while control Senders sent $3.19. However, the difference is not statistically significant at a 5% level.

Analogously, a main outcome of interest in Experiment R was whether treated Receivers, who were told that their avatar would be shown to a Sender in a Dictator Game, received more money than control Receivers, who designed their avatar before knowing how it would be used. I regress amount received by Receivers on an indicator for treatment, and control for round number. The results of the regression appear in Table 2.5: indeed, treated Receivers receive approximately 20 cents more on average, but this difference is not statistically significant at the 5% level ($p = 0.102$).

2.3.2 Treatment Effect on Avatar Ratings in Experiment R

Is there a significant treatment effect on how Receivers' avatars were perceived and rated? In addition to mapping the rating on each dimension to an integer between 1 and 6, I calculate an *index of positivity* that aggregates the eight ratings into a single value.[8] The index of positivity is a

[8] I also calculate an analogous index for the ratings dimensions, excluding *Female-Male*.

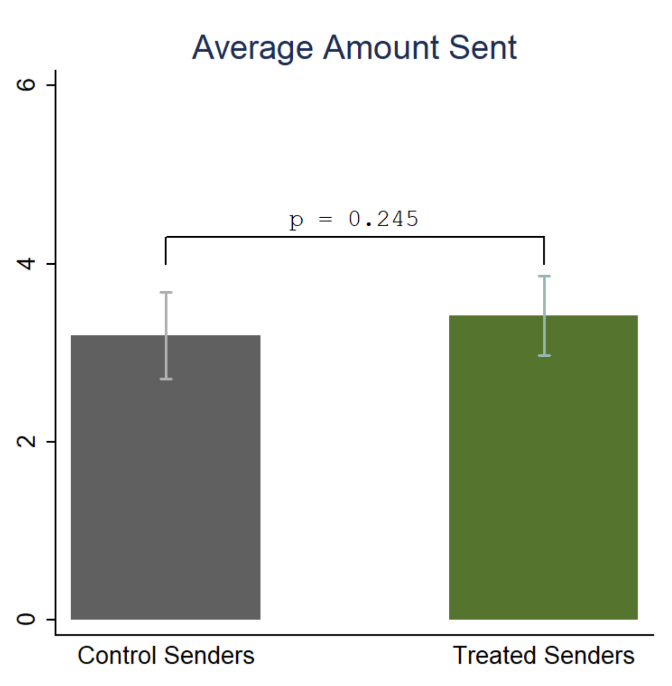


Figure 2.3: Amount Sent by Senders in Experiment S. Error bars show a 95% confidence interval. The p-value shown is from a one-sided, two-sample t-test.

	(1)
	Amount Received
isTreated_Receiver	0.194
	(0.118)
Round	-0.033
	(0.025)
Constant	3.248***
	(0.129)
Observations	1338

* p < 0.05, ** p < 0.01, *** p < 0.001

Table 2.5: Treatment Effect on Amount Received by Receiver-Creators in Experiment R. Standard errors (in parentheses) are clustered at the Sender-Rater level.

simple arithmetic mean, normalized so that the index takes values in $[0, 1]$:[9]

$$\text{Index of Positivity} = \frac{(\text{Agreeable} + \text{Attractive} + \ldots + \text{Smart}) - 8}{(48 - 8)}$$

I regress the two indices of positivity on an indicator for whether the avatar's Receiver-Creator was treated, as well as controls for the Sender-Rater's demographics and the round in which the avatar was rated. Additionally, I perform analogous regressions for the eight dimensions individually. I report the results of the former regressions in Table 2.6, and of the latter regressions in Table 2.7.

The estimations suggest a significant treatment effect. Treated Receivers' avatars were rated more positively than avatars created by Control Receivers: on average, Treated avatars were rated one point higher on one dimension than Control avatars.[10] Even when the *Female-Male* dimension is excluded from the positivity index, the positive treatment effect persists. Moreover, the estimates indicate significant, positive treatment effects for seven of eight dimensions.

It is possible, however, that the observed treatment effect could be driven by demographic differences in the Creators themselves. As a most direct example, the relatively large, positive coefficient on *Female* in column (4) could be driven by men making up a smaller proportion of Creators in the Treatment group relative to the Control group (see Table 2.4). I regress the Female dimension rating on indicators for the Creator's gender and treatment assignment, as well as the interaction, and present the results in Table 2.8.

Indeed, column (1) indicates that avatars made by male Creators are rated lower on the *Female* dimension than those made by female Creators. However, column (2) shows that the treatment effect on *Female* is robust to controlling for the Creator's gender. In fact, column (3) suggests that a significant proportion of the the treatment effect is driven by the interaction: male Creators, when they know that their avatar will be shown to a Sender, create significantly more female-presenting avatars.

[9] A positivity index of 0 maps to an avatar that is rated *Very Well* for the negative word on all eight dimensions (i.e., scores 1 on each of the eight dimensions). Conversely, a positivity index of 1 maps to an avatar that is rated *Very Well* for the positive word on all eight dimensions (i.e., scores 6 on all eight dimensions).

[10] The coefficient on the treatment indicator in column (1) of Table 2.6 is 0.025, which is the amount by which the positivity index increases when the score of a given dimension increases by 1 point.

	(1) Positivity	(2) Pos_ExclFem
isTreated_Creator	0.025***	0.015***
	(0.004)	(0.004)
Round	-0.002***	-0.002***
	(0.000)	(0.000)
Age_Rater	-0.000	-0.000
	(0.000)	(0.000)
isMale_Rater	-0.012	-0.016
	(0.007)	(0.008)
isParent_Rater	-0.001	0.001
	(0.008)	(0.009)
isWhite_Rater	-0.004	-0.006
	(0.009)	(0.010)
isLeft_Rater	0.003	0.001
	(0.009)	(0.010)
isRight_Rater	0.002	0.002
	(0.010)	(0.012)
Constant	0.606***	0.650***
	(0.016)	(0.018)
Observations	6345	6345

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.6: Treatment Effect on Perceived Positivity of Receiver-Creator Avatars in Experiment R. Standard errors (in parentheses) are clustered at the Rater level.

	(1) Attractive	(2) Female	(3) Cooperative	(4) Honest	(5) Calm	(6) Agreeable	(7) Smart	(8) Extraverted
isTreated_Creator	0.086**	0.511***	0.082*	0.085***	0.037	0.068*	0.078**	0.073*
	(0.030)	(0.050)	(0.032)	(0.026)	(0.029)	(0.029)	(0.026)	(0.030)
Round	-0.011**	0.002	-0.006	-0.003	-0.010**	-0.012***	-0.001	-0.026***
	(0.003)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Age_Rater	0.001	0.000	-0.005*	-0.002	-0.005*	-0.001	0.000	0.007***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
isMale_Rater	0.007	0.067	-0.061	-0.096	-0.092	-0.138*	-0.141*	-0.026
	(0.061)	(0.051)	(0.050)	(0.057)	(0.056)	(0.056)	(0.057)	(0.043)
isParent_Rater	0.084	-0.065	-0.063	0.056	-0.014	0.004	0.033	-0.069
	(0.067)	(0.054)	(0.052)	(0.059)	(0.059)	(0.056)	(0.061)	(0.048)
isWhite_Rater	-0.122	0.067	0.012	-0.026	0.033	-0.012	-0.117	0.025
	(0.087)	(0.065)	(0.063)	(0.064)	(0.067)	(0.070)	(0.070)	(0.055)
isLeft_Rater	0.009	0.085	0.034	-0.078	0.064	0.034	-0.022	0.009
	(0.079)	(0.066)	(0.063)	(0.069)	(0.070)	(0.068)	(0.075)	(0.060)
isRight_Rater	-0.020	0.042	0.009	-0.013	0.072	0.011	-0.030	0.027
	(0.084)	(0.072)	(0.069)	(0.079)	(0.081)	(0.079)	(0.084)	(0.068)
Constant	3.810***	2.476***	4.106***	4.390***	4.620***	4.532***	4.356***	3.938***
	(0.131)	(0.135)	(0.117)	(0.121)	(0.127)	(0.123)	(0.128)	(0.104)
Observations	6345	6345	6345	6345	6345	6345	6345	6345

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.7: Treatment Effect on Ratings of Receiver-Creator Avatars in Experiment R. Standard errors (in parentheses) are clustered at the Rater level. The word that appears at the top of each column is the word associated with the value 6; its opposite has value 1.

	(1) Female	(2) Female	(3) Female
isMale_Creator	-2.782***	-2.758***	-2.882***
	(0.05)	(0.05)	(0.07)
isTreated_Creator		0.209***	0.069
		(0.04)	(0.07)
Treated x Male			0.223**
			(0.08)
Constant	4.608***	4.483***	4.567***
	(0.04)	(0.05)	(0.06)
Observations	6669	6669	6669

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.8: Treatment and Gender Effects on Female Rating of Receiver-Creator Avatars in Experiment R. Standard errors (in parentheses) are clustered at the Rater level. The word that appears at the top of each column is the word associated with the value 6; its opposite has value 1.

2.3.3 Exploratory Analyses

Amount Received Over Time in Experiment R

Averaging across all three rounds, Treated Receivers in Experiment R earn more money than Control Receivers. Does the direction of the effect persist across all rounds? Figure 2.4 shows that the answer is yes, but that the difference is largest in the first round ($0.34), and the gap narrows in the second ($0.09) and third ($0.16) rounds.

Additionally, I consider whether the perceived positivity of a Receiver's avatar predicts the amount of money sent to the Receiver. I regress the mean amount received on the mean positivity index of their avatar, for the entire dataset (controlling for round number), and also for each round separately. The results in Table 2.9 indicate, overall, a positive relationship between the amount a Receiver receives and the positivity of their avatar. Interestingly, the direction of the effect is not consistent across rounds: the sign of the coefficient is negative in the second round.

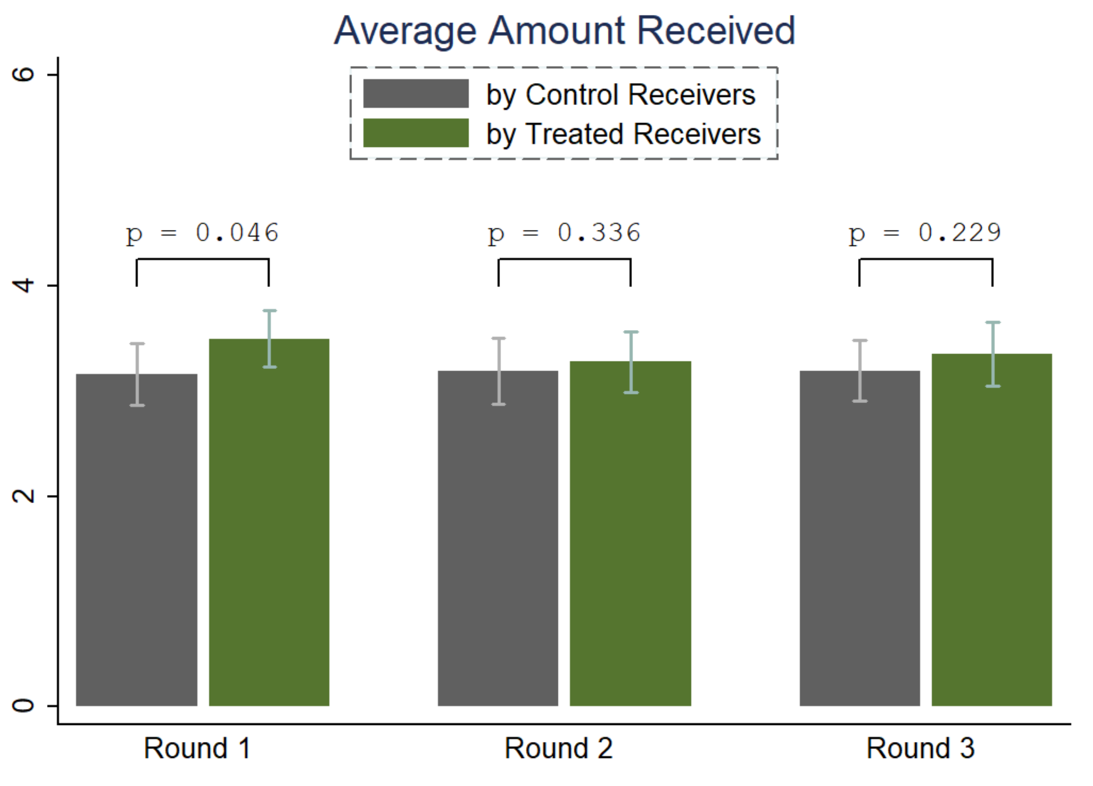


Figure 2.4: Amount Received by Receivers in Experiment R. Error bars show a 95% confidence interval. The *p*-values shown are from one-sided, two-sample *t*-tests.

	Round 1	Round 2	Round 3	All Rounds
Positivity_Mean	1.495	-0.801	0.571	0.449
	(0.95)	(1.04)	(1.04)	(0.60)
Round				-0.038
				(0.02)
Constant	2.465***	3.700***	2.938***	3.098***
	(0.57)	(0.61)	(0.61)	(0.36)
Observations	446	446	446	1338

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.9: Effect of Receiver Avatar's Positivity on Amount Received by Receiver in Experiment R. Standard errors are clustered at the Sender-Rater level.

Demographics of Senders in Experiment S

In Experiment S, I find a small, positive treatment effect on amount sent. Is the direction of the treatment effect robust to controls for Sender demographics? I regress amount sent on an indicator for whether the Sender was treated, and control for the demographic characteristics of the Sender as well as interactions of the treatment dummy and the demographic characteristics. Column (2) in Table 2.10 shows that the direction of the treatment effect stays positive when controlling for demographics and reveals a significant, positive impact of being a parent on amount sent to the Receiver. Moreover, column (3) of the same Table indicates significant interactions of the treatment with subjects' political leanings.

2.4 Conclusions

As more individuals represent themselves with avatars in digital settings, it is imperative to understand how avatars induce self image and social image considerations in users. In Experiment S, the data suggest that seeing one's avatar does not induce significantly higher prosociality than when one doesn't see one's own avatar. On its face, this result does not align with the result in [14], in which subjects who saw themselves chose to shock an anonymous participant significantly less often than subjects who saw a neutral stimulus (or not stimulus at all). However, another distinction between Experiment S and the [14] experiment is in the task: it is reasonable to think that dividing a $10 endowment (as in Experiment S) is a less consequential decision than administering a painful electric shock to somebody else. The lack of a significant treatment effect in Experiment S may also be driven by this difference.

In Experiment R, the data suggest that individuals do design their avatars differently when they are aware of how their avatar will be used. The treatment effect moves avatars toward more 'positive' traits—in particular, toward being more attractive, smart, and honest—and induces treated, male Creators to make avatars that look more female. Moreover, there is an economically meaningful treatment effect: treated Receiver-Creators receive more money than do control Receiver-

	(1) Amount Sent	(2) Amount Sent	(3) Amount Sent
isTreated_Sender	0.23	0.10	-1.79
	(0.33)	(0.34)	(0.95)
Age		-0.01	-0.00
		(0.01)	(0.01)
isMale		-0.30	-0.63
		(0.36)	(0.53)
isParent		1.34***	1.05*
		(0.37)	(0.53)
isWhite		0.33	0.43
		(0.36)	(0.48)
isLeft		0.18	-0.86
		(0.41)	(0.60)
isRight		-0.48	-1.89**
		(0.48)	(0.60)
Treated x Male			0.44
			(0.71)
Treated x Parent			0.35
			(0.71)
Treated x White			-0.30
			(0.67)
Treated x Left			1.93*
			(0.81)
Treated x Right			2.72**
			(0.90)
Constant	3.19***	2.87***	3.92***
	(0.24)	(0.70)	(0.85)
Observations	243	225	225

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 2.10: Amount Sent by Senders in Experiment S. Standard errors are shown in parentheses.

Creators. However, the former effect doesn't explain the latter: variation in money received is not explained by variation in rating on any particular dimension. The mechanism that drives the increased amount received by treated Receiver-Creators remains an open question for future work.

Chapter 3: Minority Turnout and Representation under Cumulative Voting. An Experiment.

3.1 Introduction

The fragility of American democracy, rooted historically in slavery, manifests itself in persistent efforts to disenfranchise racial and linguistic minorities, Black Americans first and foremost. Almost 60 years after the Voting Rights Act (VRA), the disputes we continue to witness are reminders of the heightened importance of voters' participation. In 2012, the Pew Research Center concluded: "The Growing Electoral Clout of Blacks Is Driven by Turnout".[1] The date is not coincidental: 2012 was the election year for President Obama's second term, when for the first time, Black turnout was higher than White turnout.[2]

Guaranteeing high electoral participation by minorities requires rules about fair and equal access to voting. But that is not enough: as the surge in Black political engagement during the Obama years shows, it also requires giving minorities the realistic chance of a desired outcome.[3] America's majoritarian electoral system makes this difficult. Without resorting to proportional representation, the courts have mandated modifications to electoral rules in jurisdictions where majoritarian systems effectively disenfranchise the minority. The traditional remedy has been the design of single member districts in which the minority constitutes a majority of the electorate. In 1993, however, the Supreme Court judged unconstitutional districting plans driven by considerations of race (*Shaw v. Reno*), and such districts have since had a troubled legal history. Our focus

[1] [59]

[2] According to the US Census Bureau, Black non-Hispanic turnout increased from from 60% to 65% and then to 67% from 2004 to 2008 to 2012, while, over the same period, the turnout of non-Hispanic Whites went from 67% to 66% and then 64%.

[3] For a brief panoramic summary of voter suppression in the US and the role of the VRA, see [61]; for a discussion of theories of representation and empowerment, and for empirical evidence, see [62].

is on a different alternative proposed by the courts: Cumulative Voting (CV), a solution built not on controlling district borders but on the voting rule itself.

CV applies to elections in multi-member districts. The core idea is to allow voters to vary the number of votes cast for each candidate. Under CV, each voter has as many votes as there are open seats, and the candidates with more votes win, as under simple plurality. However, each voter is allowed to distribute the votes freely among any number of candidates. CV treats every voter equally; yet, a cohesive minority can ensure itself some victories by cumulating its vote. CV is not conditional on fixed patterns of geographical segregation and thus does not require adjustments as social and political conditions change. As a result, it is not subject to the type of litigation that has weakened the VRA. In addition, although CV can deliver semi-proportional outcomes, it does so through a relative minor modification of a majoritarian system. The US, the UK, Scotland, countries with long-held skepticism of proportional representation have a history of accepting CV.[4]

This paper contributes to the debate on minority representation by analyzing, both theoretically and experimentally, CV's potential to increase the relative voting participation of the minority and its share of seats in elected bodies. Both effects have been observed in actual applications, but evaluating historical evidence is complicated by the non-random adoption of CV. The existing evidence thus must be accompanied by experimental testing. Our main conclusion is that, in the controlled environment of the lab, CV's predicted outcomes are realized. Across all experimental parametrizations, the relative participation in voting of the minority group increases and so does its share of electoral victories. Part of the impact on minority victories stems immediately from the allowed cumulation of votes: to prevent spreading votes too thinly, the majority must limit the number of candidates it fields, leaving openings for minority candidates. But another important contributor to the minority's success is that CV increases the differential turnout of the minority, relative to the majority. Both theoretically and in the lab, the improved prospects brought by CV work to increase the fraction of minority voters among those who turn out. As in the Obama

[4]Similarly to CV, Limited Vote (LV) also results in semi-proportional outcomes. Under LV, voters have fewer votes than the number of candidates and cast one vote per chosen candidate. LV is considered simpler than CV but less reliable in generating minority representation. See for example [63]. For a broad discussion of alternative rules and proportional representation, see [64].

election or the Pew Research report summary title mentioned above, the realistic promise of representation encourages political participation.

CV was used for more than 100 years, from 1870 to 1980, to elect representatives to the Illinois State House and is the rule now in the election of local commissions in tens of local jurisdictions. Empirical case studies have been promising. In Alamogordo, New Mexico, Latinos amounted to 20% of the electorate but had long been unrepresented in the City Commission. In 1987, after the adoption of CV, the City Commission welcomed its first Latino representative in twenty years. In Amarillo, Texas, minorities made up 24% of the electorate, but lacked representation on the school board. In 2000, after the adoption of CV, Amarillo welcomed its first Black and first Latina school board representatives. In Chilton County, Alabama, the Black community (11% of the electorate) struggled for years with unpaved roads. In 1988, after the adoption of CV, the first Black county commission member was elected since Reconstruction. The roads were finally paved.[5] Empirically, then, CV correlates with an increase in the number of elected minority representatives ([70]; [65]), and in the public goods provided to minority communities ([68]). In addition, its use appears to increase minority participation in the political system, and in a study analyzing the impact on local US jurisdictions, CV is associated with an increase of approximately 5 percentage points in overall turnout ([71]).

As inspiring as these results are, when implemented, CV typically follows voting rights litigation, indicating heightened sensitivity to minority representation and stronger minority involvement. Understanding the specific role of the voting rule is helped by complementing the historical experiences with the study of a move to CV under the controlled conditions of the lab. We run different experimental treatments, comparing standard bloc voting (one-vote-per-open-seat) and CV, and varying both the number of seats and the relative size of the minority. Because our more innovative contribution concerns turnout, we focus on a voter's incentive to overcome obstacles to voting. We run a canonical costly voting experiment where payoffs depend on one's own group

[5] See [65] for a short history of CV. Other useful sources are [66], [67], [68]. For a strong defense of CV, see [69]. Updated information on the current use of CV is reported in fairvote.org. Outside local politics, CV is used to elect corporate boards in approximately 10% of S&P 500 companies, again with the goal of protecting minority representation.

achieving electoral success but voting is individually costly ([72]; [73]). Do participants, particularly participants on the minority side, overcome those costs more often when votes can be cumulated?

Empirically, unless complemented by sophisticated formulations of bounded rationality, costly voting models fail to predict the level of turnout in large elections. However, in studies whether of historical or experimental data, their comparative predictions have fared better: turnout is predicted to increase when elections are closer, when they are more salient, when voting costs are lower, when the electorate is smaller (for example, [74]; [73]). It is this type of comparative effect that interests us: when the voting rule changes to CV, are minority voters more represented among overall voters? Precise theoretical predictions depend on complex calculations of pivotality, but the logic underlying the results is in fact much simpler: when votes can be cumulated, and only when votes are cumulated, the minority can win seats even when if realized turnout rates are similar between minority and majority voters. In equilibrium, such prospects encourage minority participation: CV always increases the differential turnout of the minority, relative to the majority.

According to the theory, the change in differential turnout is the result of a sharp decline in majority voting participation, together with either a mild increase or a a minor decline for the minority. In our experiment, although the impact on differential turnout is realized and is robust, turnout under CV remains higher than expected for both groups. We read this evidence as suggesting that the subjects are responding to the increased competitiveness introduced by CV and the minority's heightened prospects more closely than to the pivotality calculations that underlie the exact theoretical prediction.

To our knowledge, there is no existing theoretical or experimental study of turnout under CV. Previous laboratory experiments on CV ([75] and [76]) focus on the coordination problem the voting rule poses and neglect the impact of the voting rule on voters' participation decision. We take the opposite approach. We focus on voters' turnout decisions and assume that the coordination problem is addressed by the parties' leadership, and addressed primarily through the leaders' choice of the number of party candidates.

We make this assumption because it mirrors our reading of CV's historical experiences. For example, [66] is a very lively study of CV in Victorian England, an interesting environment for its experimental spirit, the richness of cases, and the availability of historical documents. Focusing exactly on the strategic problems posed by CV, the authors find: "a willing demand for party organization from voters, as much as a willing supply of it from the parties themselves." Strategic mistakes were made, typically in the form of over-nominations by the majority party, but their responsibility was attributed to party leaders, and quickly corrected. Consider the following exchange from 1884 Parliamentary hearings on an election run with CV (cited by [66], p. 911):

Mr. Courtney: If a party ran too many candidates it might not gain its due proportion of power.
Mr. Sanford: Quite so. That is its own fault.[6]

Similar sentiments recur in other episodes, whether the majority and minority identities are party-based, as in Victorian England, or correspond to racial or linguistic division, as described, for example, in [68]'s detailed chronicle of the first adoption of CV in Chilton County, Alabama, following VRA litigation. Because they are so costly, nomination mistakes are corrected rapidly, and granting party leaders their optimal choice of candidates seems a good working assumption.[7] Note also that a common finding in the literature is that nomination mistakes are more common on the majority side, for whom the need to concentrate votes is less obvious. If our analysis underestimates the parties' difficulties in coordinating votes, it is likely to also underestimate the extent to which the minority benefits from CV.

Because CV is an example of "semi-proportional" voting rules—rules whose results approach proportional representation without imposing proportionality—parallel to our work are the experiments in [79] and [80], which compare turnout under single-winner majoritarian and proportional elections. However, although CV leads to quasi-proportional outcomes, the turnout decision is

[6][77]. School board elections attracted much attention because they decided religious education in schools, and in particular the inclusion or not of Catholicism.

[7]When CV becomes established, if anything the often voiced concern is the possibility of collusion between party leaders, reducing voters' choices, as was remarked for example during the long experience with CV in the Illinois State House elections ([78]).

quite different: under proportional representation, the value of a marginal single vote is proportional to the change in the party's vote share, and pivotality, in its usual sense, is moot.[8] With CV, instead, pivotality continues to drive turnout decisions. The difference, relative to majoritarian voting, is that the possibility of cumulating votes implies richer pivotality calculations. This said, the conclusions are similar: both Herrera et al. and Kartal find that proportional representation increases the turnout rate of the minority relative to the majority's, and the minority's expected share of power. The same results hold under CV.

Currently, CV is limited to elections of local committees, and it is natural to ask whether there are realistic chances of applications at a higher level. One possible reason for skepticism is that CV applies to multi-member district elections and recourse to such elections is relatively infrequent, not least because of its long history of legal challenges, exactly on the grounds of discrimination against racial minorities. The problem, however, is the combination of multi-member districts and bloc voting: in the absence of cumulation, a group that has a minority position in all districts can indeed potentially win no representation at all. CV, on the contrary, favors minority victories, in line with the Courts' recommendation of its adoption following voting rights litigation. At the moment, ten US states use multi-member district elections to elect at least one of their state chambers.[9] Such elections could in principle adopt CV and indeed there have been repeated calls to that effect.[10] With broader stakes, the focus would presumably be less on the representation of racial and linguistic minorities and more on parties' dynamics. The need to understand better the effects of the voting rule remains.

The paper proceeds as follows. The next section describes the basic model in the absence of voting costs and compares equilibrium minority victories under one-vote-per-seat and CV. Section 3.3 discusses theoretical predictions under voting costs. Section 3.4 describes the experiment, and Section 3.5 discusses the experimental results. Section 3.6 concludes. Additional theoretical

[8] Indeed [79] comment on the similarity in turnout decisions between proportional voting models and non-instrumental models of voting.

[9] Arizona, Idaho, Maryland, New Hampshire, New Jersey, North Dakota, South Dakota, Vermont, Washington, and West Virginia. Ten more states explicitly allow multi-member districts by law.

[10] See for example [81], for a return to CV in Illinois state elections, or the advocacy of Fairvote.org.

material is left to the Appendix. An online Appendix reports supplementary empirical results, as well as a copy of the experimental instructions.

3.2 Base Model

An electorate of N potential voters selects $K > 1$ representatives for a commission. All positions are identical, and all are simultaneously decided in the election. The N voters are divided into two parties: M, the majority party with M members, and m, the minority party with $m < M$ members, where $M + m = N$. Parties are led by party leaders whose role is to propose the party's list of candidates.

Within each party, all potential candidates are identical, and party leaders and voters share the same objective: to maximize the number of positions won by their party. The utility derived from one's party winning k positions is $u(k)$, increasing in k. We denote by V the value of controlling all positions and assume $u(k) = (k/K)V$. Linearity captures the focus on the number of positions and simplifies both the lab implementation and the theoretical analysis. But we adopt it on substantive grounds as well: any "place at the table" has value. The assumption mirrors an exercise of committee power that is proportional to the number of seats a party has won.

Each voter has K votes, and the K candidates with most votes are elected. If there are ties, after the highest voted candidates are elected, the remaining open positions are filled by selecting winners randomly among the tied candidates. We call x_p the profiles of votes cast by members of party p, where x_{ip}^k is the number of votes cast by voter $i \in p$ for candidate k, and x_{ip} the vector of all votes cast by i.

We study two electoral systems, *multi-seat plurality* (MP) and *cumulative voting* (CV). MP corresponds to standard bloc voting in multi-member districts: under MP, each voter casts at most one vote for each candidate: $x_{ip}^k \in \{0, 1\}$ for all i, k, and p, and each party nominates K candidates.[11] Under CV, each voter can distribute the K votes in any manner the voter desires, as long as the overall budget of K votes is satisfied: $\sum_k x_{ip}^k \leq K$. The possibility of cumulating votes creates

[11]MP is used by all but two of the US states electing their legislatures from multi-member districts.

a coordination problem that party leaders help address by selecting the number of candidates, G for the majority party, and g for the minority party. In line with historical experience[12], we allow for fractional votes, but voters, candidates, and positions are constrained to be integers.

The game has two stages. In the first stage, party leaders announce the party list; in the second stage, voters distribute their votes over the party candidates. We focus on equilibria in weakly undominated strategies where voters cast all their votes and cast votes on their party's candidates only. Under MP, each voter casts one vote for each party candidate. Under CV, the equilibrium is a pair of vote profiles $\{x_{\mathsf{M}}(G, g), x_{\mathsf{m}}(G, g)\}$ and a pair of party lists $\{G(x_{\mathsf{M}}, x_{\mathsf{m}}), g(x_{\mathsf{M}}, x_{\mathsf{m}})\}$ such that each party member's votes maximize the number of seats won by the party, given the parties' lists and the other voters' voting choices, and each party list maximizes the number of seats won by the party, given the opposite party's list and all voters' voting strategies.

Although our main focus is on turnout, and thus the theory will require some positive costs of voting, it is helpful to begin by understanding the functioning of the two voting rules without the complication of voting costs.

3.2.1 Minority representation without voting costs

With no reasons to abstain and no leeway in distributing votes, under MP, party M wins all seats: each M candidate receives M votes, and each m candidate receives $m < M$ votes.

CV grants the minority the possibility of winning some seats. Suppose for example that all voters in party m concentrate all their votes on a single candidate, who thus receives mK votes. The minority wins one seat if its candidate beats the Kth weakest majority candidate, that is, the majority party's candidate who ranks Kth in terms of votes received. If the majority nominates fewer than K candidates, the minority candidate is elected. If the majority targets all positions and nominates K candidates, the weakest majority candidate will have most votes when the MK total majority votes are distributed equally among the K majority candidates, and each receives $MK/K = M$ votes. Hence the minority can *guarantee* itself a seat if $mK > M$, or $m > M/K$.

[12] For example, half votes were allowed in the Illinois State House; half, third, and quarter votes are allowed in the Peoria, IL elections.

This ratio, known in the literature as the *threshold of exclusion*, is a fraction of M: for example, a minority that is half the size of the majority can guarantee itself a seat if the number of open seats is three or more.

Academics and lawyers have extended this logic to a handy formula that delivers a party's guaranteed number of seats for each $\{m, M, K\}$.[13] On the face of it, the formula does not address what we are really interested in: not how many seats can the minority make sure to win, but how many seats will it win when both parties play their optimal strategies. Yet the answer the formula yields can be grounded in a strategic analysis. In line with our focus on the coordinating role of the party leaders, we call *party-optimal* those equilibria that for each party maximize the number of seats won. We denote by z the number of seats won by party m. As we prove in the Appendix:

Proposition 1. *In the absence of voting costs, in all party-optimal equilibria of the CV voting game: (i) for all $m < M/K$, the minority never wins any seat; (ii) for all $m \geq M/K$:*

$$z = \left\lfloor \frac{Km+m}{M+m} \right\rfloor \qquad \textit{if } \frac{Km+m}{M+m} \notin \mathbb{Z}$$

$$z = \begin{cases} \frac{Km+m}{M+m} - 1 & \text{with prob } m/(m+M) \\ \frac{Km+m}{M+m} & \text{with prob } M/(m+M) \end{cases} \qquad \textit{if } \frac{Km+m}{M+m} \in \mathbb{Z}$$

Given M, m, and K,the proposition yields the equilibrium number of minority seats. Suppose, for example, $m = M/2$. Then $z = 1$ if $K = 4$; $z = 2$ if $K = 6$; and $z = 1$ with probability $1/3$ or $z = 2$ with probability $2/3$ if $K = 5$.

What makes the result powerful is the unique equilibrium prediction on the number of minority seats. As we discuss in more detail in the Appendix, for given M, m, and K, the game admits a large number of party-optimal equilibria. And yet the multiplicity is irrelevant to the outcome: *all*

[13]The formula is so widely known and used that CV-calculators can be found online. See, for example, https://www.lawjock.com/tools/cumulative-voting-calculator/, or Wikipedia https://en.wikipedia.org/wiki/Cumulative_voting. Early influential references in political science are [82], [83], [78], [84], and [85].

party-optimal equilibria must yield the same number of minority victories.

The result follows from two main reasons. First, because voters and leaders share a common goal, party-optimal equilibria correspond to the equilibria of a two-player game where the two party leaders directly control the distributions of the votes over the party candidates. Second, the linearity of the objective function, $u(k)$, renders the game constant-sum. As a result, all party-optimal equilibria must result in maximin payoffs, the payoffs the two parties can guarantee themselves. Extending the reasoning described earlier then yields the proposition.

Note an immediate corollary that will shape intuition for what follows. There always exists an equilibrium where $g = \lfloor \frac{Km+m}{M+m} \rfloor$, $G = K - \lceil \frac{Km+m}{M+m} - 1 \rceil$, and $x_{im}^k = K/g$, $x_{iM}^k = K/G$: all voters spread their votes equally over their party's candidates, and the two parties nominate just enough candidates to fill all open positions if $\frac{Km+m}{M+m} \notin \mathbb{Z}$, or exceed the number of positions by 1 if $\frac{Km+m}{M+m} \in \mathbb{Z}$.

3.3 Voting costs

Suppose now that each voter i faces a cost of voting c_i, drawn randomly and independently across voters from a common distribution $F(c)$ everywhere continuous and atomless over support $[\underline{c}, \overline{c}]$, with $\underline{c} \geq 0$. Realized costs are private information, but the distribution $F(c)$ is common knowledge and does not depend on party affiliation. The cost c_i represents the cost of going to the polls and is independent of the number of votes cast. A voter i whose party wins k positions has utility $U_i(k)$, given by:

$$U_i(k) = \begin{cases} u(k) - c_i & \text{if voter } i \text{ voted} \\ u(k) & \text{if voter } i \text{ abstained} \end{cases}$$

3.3.1 Multi-winner plurality (MP)

Under MP, voters who have turned out cast a single vote for each of the party's K candidates. Although multiple positions are in play, the analysis mirrors closely the standard approach to costly voting in single winner elections.[14] Following [72], and the subsequent literature, we focus on semi-symmetric Bayesian equilibria in threshold strategies: there exist cost thresholds c_{M} and c_{m} such that any voter i in party M (m) turns out to vote if $c_i < c_{\mathsf{M}}$ ($c_i < c_{\mathsf{m}}$) and abstains if $c_i > c_{\mathsf{M}}$ ($c_i > c_{\mathsf{m}}$).

Call S_p the number of voters who turn out for party p. Each M candidate receives S_{M} votes, and each m candidate receives S_{m} votes. Thus only three outcomes are possible: either $S_{\mathsf{M}} > S_{\mathsf{m}}$, and all K positions are won by M candidates; or $S_{\mathsf{M}} < S_{\mathsf{m}}$, and all K positions are won by m candidates; or $S_{\mathsf{M}} = S_{\mathsf{m}}$, and all K positions are tied, with K majority and K minority candidates all having the same number of votes. Under a tie, the K winners are chosen randomly among all tied candidates. We denote by Eu_T^{MP} the expected utility gain from winning seats under MP in case of a tie. Then:

$$Eu_T^{MP} = \sum_{k=0}^{K} \frac{\binom{K}{k}\binom{K}{K-k}}{\binom{2K}{K}} u\left(k\right) = V/2$$

where the second equality follows from $u(k) = (k/K)V$.

As in single-winner elections, a voter from party p facing opposite party p' must weigh her cost of voting against the expected utility gain from influencing the outcome. Denoting by S_{-ip} the number of voters who turn out in party p ignoring i, voter i can influence the outcome either by breaking ties (when $S_{-ip} = S_{p'}$; an event whose probability we denote by π_p^T) or by making ties (when $S_{-ip} = S_{p'} - 1$, with probability π_p^{T-1}). Thus if the thresholds $\{c_{\mathsf{M}}, c_{\mathsf{m}}\}$ are interior, they solve the system of equations:

$$c_{\mathsf{m}} = [u(K) - Eu_T^{MP}]\pi_{\mathsf{m}}^T(c_{\mathsf{M}}, c_{\mathsf{m}}) + [Eu_T^{MP} - u(0)]\pi_{\mathsf{m}}^{T-1}(c_{\mathsf{M}}, c_{\mathsf{m}}) \tag{3.1}$$

$$c_{\mathsf{M}} = [u(K) - Eu_T^{MP}]\pi_{\mathsf{M}}^T(c_{\mathsf{M}}, c_{\mathsf{m}}) + [Eu_T^{MP} - u(0)]\pi_{\mathsf{M}}^{T-1}(c_{\mathsf{M}}, c_{\mathsf{m}}) \tag{3.2}$$

[14][86] study costly voting with multiple candidates but a single winner. Our model is closer to the traditional two-candidate, one-winner set-up, with each party list being the parallel to the party candidate.

or:

$$c_{\mathsf{m}} = (V/2)\pi_{\mathsf{m}}(c_{\mathsf{M}}, c_{\mathsf{m}}) \tag{3.3}$$

$$c_{\mathsf{M}} = (V/2)\pi_{\mathsf{M}}(c_{\mathsf{M}}, c_{\mathsf{m}}) \tag{3.4}$$

where $\pi_p \equiv \pi_p^T + \pi_p^{T-1}$ is the pivotal probability for a voter of party p.

The linearity of the utility function implies that the equilibrium equations (3.3) and (3.4) do not depend on K. The problem is then formally identical to the classic costly voting problem with a single winner and two alternatives. It is well-known, and we leave the expressions for the pivot probabilities to the Appendix. Given equilibrium $\{c_{\mathsf{m}}, c_{\mathsf{M}}\}$, we can derive the probabilities of winning different number of positions. The derivation is straightforward, and again is left to the Appendix.

3.3.2 Cumulative voting (CV)

With voting costs, the interests of voters and party leaders need not coincide any longer. The game now has $M + m + 2$ players and three stages: a nomination stage, when the two leaders choose the number of candidates; a turnout stage, when, after observing privately the realization of the voting cost, each voter decides whether or not to vote; and finally a voting stage, when voters at the polls choose how to cast their votes.[15]

We focus on pure strategy semi-symmetric perfect Bayesian equilibria such that within each party, all voters follow the same strategy. We denote by x_{-ip} the profile of votes cast by voters other than i who have turned out and belong to p. The equilibrium is a pair of party lists $\{g, G\}$, a pair of cost thresholds $\{c_{\mathsf{M}}, c_{\mathsf{m}}\}$, and a pair of voting profiles $\{x_{\mathsf{M}}, x_{\mathsf{m}}\}$ such that: (i) at the voting stage, voter i in party p who has gone to the polls sets $x_{ip}(G, g, c_M, c_m, x_{-ip}, x_{p'})$ so as to maximize the expected number of positions won by p, and in equilibrium $x_{ip}^k = x_p^k$ for all i and $k \in p$; (ii) at the turnout stage, all $i \in p$ with $c_i < c_p(G, g, c_{p'}, x_M, x_m)$ strictly prefer to vote, and all $i \in p$

[15]Because party leaders influence turnout through the number of candidates, note the connection to models of leaders' enforced social norms in voting ([87]).

with $c_i > c_p(G, g, c_{p'}, x_M, x_m)$ strictly prefer to abstain; and (iii) at the nomination stage, the two party leaders set $g(G, x_M, x_m, c_M, c_m)$ and $G(g, x_M, x_m, c_M, c_m)$ so as to maximize their party's expected number of positions. The term "equilibrium" in what follows, refers to such equilibria.

For any positive turnout, if $g < K$, party M is guaranteed $\min[G, K - g]$ seats, and similarly, if $G < K$, party m is guaranteed $\min[g, K-G]$ seats. The positions contested are $\max[0, g+G-K]$.[16]

The voters' turnout decision complicates greatly the characterization of the equilibrium. Intuitively, there are three main reasons. First, turnout is stochastic and in evaluating how to distribute votes, each voter needs to account for the probability of different turnout rates, both among voters of her own party and among opponents. Second, the number of candidates nominated by the party leaders will affect not only the distribution of votes among voters at the poll, but the decision to turnout itself—the cost thresholds. And because such influence is mediated by the votes profiles, the link in general is complex. Third, the multiplicity of equilibria noted in the absence of voting costs continues to exist when voting is costly. And because the game cannot be assimilated to a zero-sum two-player game any longer, the multiplicity of equilibria will in general translate into multiplicity in outcomes.

This said, two limited results must hold and will help the experimental design. We summarize them in one proposition. Recall that an equilibrium is strict when deviation implies a non-zero loss.

Proposition 2. *(i) There exists no strict equilibrium with $g + G < K + 1$. (ii) If there exists an equilibrium with $g + G = K + 1$, then the equilibrium has equal spreading of votes: $x^k_{i,\mathsf{m}} = K/g$ and $x^k_{i,\mathsf{M}} = K/G$.*

Proof. (i) If $g + G \leq K$, there are no contested seats. With no contested seats, no voter with positive voting costs goes to the polls. Because non-contested positions are ensured, for given opposite party list, increasing the number of party candidates cannot cost any seat. (ii) If $g + G =$

[16] We assume that non-contested positions are assigned to candidates nominated by the parties even in the absence of voters' turnout.

$K+1$, there is a single contested seat, and the competition between the two parties is over protecting the least voted of their respective candidates. We are focusing on semi-symmetric equilibria, where therefore $x^k_{ip} = x^k_p$ for all $i \in p$: in equilibrium all voters in p cast the same number of votes on party candidate k. Can there be an equilibrium where there exist two candidates from the same party, k and k', such that $x^k_p > x^{k'}_p$? All party candidates with the exception of the single least voted candidate are guaranteed election. Thus, if there exist k and k' such that $\sum_{-i\in p} x^k_{-i,p} > \sum_{-i\in p} x^{k'}_{-i,p}$, voter i in party p benefits from deviating. Rather than casting $x^k_{ip} > x^{k'}_{ip}$ and reinforcing the difference in votes across the candidates, i should counter it, and cast her votes so as to maximize the votes total of the least voted of the party's candidates: $Max_{\{x_{ip}\}}(Min_k(\sum_{i\in p} x^k_{ip}))$. Because by construction such a deviation increases the votes of the least voted candidate, it cannot cost any seat and is profitable if the voter is pivotal. Hence in any semi-symmetric equilibrium, if $g+G = K+1$, it must be that $x^k_p = x^{k'}_p$ for all k, k', or $x^k_{i,\mathsf{m}} = K/g$ *and* $x^k_{i,\mathsf{M}} = K/G$. □

In what follows, we identify and use as theoretical references equilibria with equal spreading of votes. We do so when $g + G = K + 1$, but also when $g + G > K + 1$. Distributing votes equally is an easy default for the voters, but we focus on such equilibria for two additional reasons. First, observers have documented that equal spreading of votes over all party candidates was the norm in the Illinois State House ([78]; [88]). With more than a century of experience, it seems plausible that such behavior condensed CV's lessons when parties play their coordinating roles. Second, the explicit constraint that votes must be spread equally is part of a modified CV rule ("Equal and even CV") applied in elections in Peoria, IL and at times proposed, because of its simplicity, as a possible model for wider adoption.[17]

To characterize equilibria for the experimental parametrizations, we begin by discussing the derivation of the equilibrium cost thresholds, and thus turnout, and showing how such derivation differs from the usual approach. For given g and G, equilibrium cost thresholds continue to trade off costs of voting and expected utility gains from influencing the election. As before, a voter may break an existing tie or cause a tie, but if the party's candidates are fewer than the number

[17] See, for example, the discussion by fairvote.org in http://archive.fairvote.org/factshts/comparis.htm

of seats, by casting more than a single vote on each, the voter may also move the outcome from a loss to a win of all contested positions. Consider the problem for $i \in \mathsf{M}$. By voting, i breaks a tie if $(K/G)S_{\mathsf{M-i}} = (K/g)S_{\mathsf{m}}$, or $S_{\mathsf{M-i}} = S_{\mathsf{m}}(G/g)$; i causes a tie if $(K/G)(S_{\mathsf{M-i}} + 1) = (K/g)S_{\mathsf{m}}$, or $S_{\mathsf{M-i}} = S_{\mathsf{m}}(G/g) - 1$. In addition, voter i can shift M from losing to winning all contested positions if both $(K/G)S_{\mathsf{M-i}} < (K/g)S_{\mathsf{m}}$ and $(K/G)(S_{\mathsf{M-i}}+1) > (K/g)S_{\mathsf{m}}$, or $S_{\mathsf{M-i}} \in (S_{\mathsf{m}}(G/g) - 1, S_{\mathsf{m}}(G/g))$. Denoting by π_p^W the probability that the votes of a member of party p move party p from losing to winning all contested positions, if c_{M} is interior, and $G + g > K$, c_{M} must solve:

$$c_{\mathsf{M}} = [u(G) - Eu_{T,\mathsf{M}}^{CV}(G,g)]\pi_{\mathsf{M}}^T + [Eu_{T,\mathsf{M}}^{CV}(G,g) - u(K-g)]\pi_{\mathsf{M}}^{T-1} + [u(G) - u(K-g)]\pi_{\mathsf{M}}^W$$

where:[18]

$$Eu_{T,\mathsf{M}}^{CV}(G,g) = \sum_{x=0}^{G} u(x)\binom{G}{x}\binom{g}{K-x}\Big/\binom{G+g}{K} = \frac{G}{g+G}V.$$

Or:

$$c_{\mathsf{M}} = \frac{V(g+G-K)}{K}\left[\frac{G}{g+G}\pi_{\mathsf{M}}^T + \frac{g}{g+G}\pi_{\mathsf{M}}^{T-1} + \pi_{\mathsf{M}}^W\right] \tag{3.5}$$

The problem is analogous for minority voters. The equilibrium condition for an interior threshold c_{m} is:

$$c_{\mathsf{m}} = \frac{V(g+G-K)}{K}\left[\frac{g}{g+G}\pi_{\mathsf{m}}^T + \frac{G}{g+G}\pi_{\mathsf{m}}^{T-1} + \pi_{\mathsf{m}}^W\right] \tag{3.6}$$

The pivot probabilities and the probabilities of winning different numbers of position in case of ties can be derived as under MP, taking into account that the number of candidates, in each party, may differ from the number of seats. We leave them to the appendix.

Given (3.5) and (3.6), and positing $x_{i\mathsf{m}}^k = K/g$, $x_{i\mathsf{M}}^k = K/G$, we can find party leaders' optimal choice of G and g. Given G, g, c_{M}, c_{m}, and the conjecture that all other voters spread votes equally, we can verify that equal spreading is a best response for a voter at the polls. Hence the solution is an equilibrium.

[18] Throughout the paper, we use the convention $\binom{r}{y} = 0$ if $y > r$.

3.3.3 Equilibria for the experimental parametrizations

Figure 3.1 shows the equilibrium cost thresholds in the two parties, $\{c_{\mathrm{m}}, c_{\mathrm{M}}\}$, and the expected fraction of seats won by the minority under the two voting systems for a set of parameters that include those used in the experiment. The distribution $F(c)$ is Uniform over $[0, 1]$, and thus the cost thresholds are equal to the two parties' turnout rates. The first column corresponds to MP, the second and third to CV (for $K = 2$ and $K = 4$, respectively; recall that K does not affect outcomes under MP). The number of candidates, G and g, equals K for MP and is set at equilibrium value for CV. In each panel, the horizontal axis corresponds to different values of M, while upper and lower panels refer to different relative sizes of the two parties.[19]

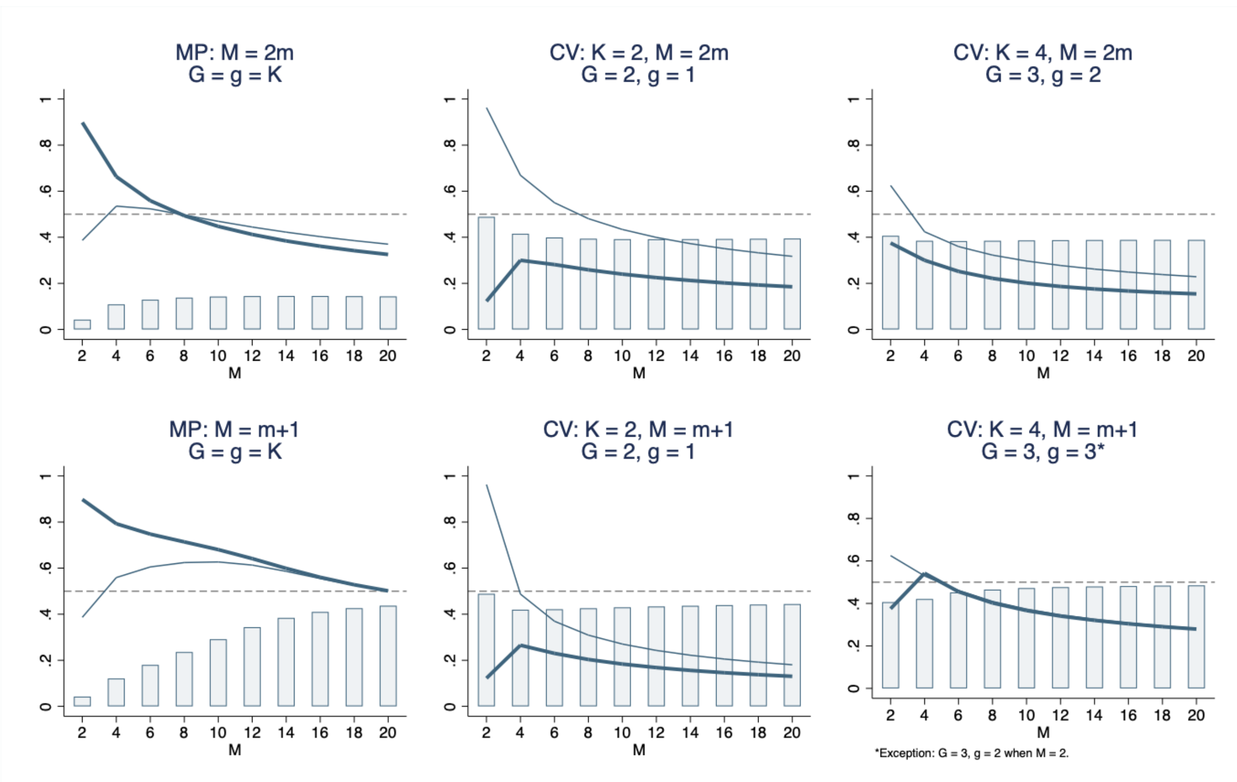


Figure 3.1: *Expected turnout rates and share of minority seats, MP and CV.* The thick lines correspond to c_{M}, the thin lines to c_{m}; the bars correspond to the expected share of minority seats. F is uniform over $[0, 1]$; $V = 4$.

[19] We found a unique equilibrium in all cases. We discuss in the Appendix the surprising lack of a consistent underdog effect ($c_{\mathrm{m}} > c_{\mathrm{M}}$) in the MP model. For both MP and CV, raising K to 6 does not change the qualitative results.

The figure highlights two main regularities. First, the differential between the minority's and the majority's turnout rates, $c_{\mathsf{m}} - c_{\mathsf{M}}$, is consistently higher under CV than under MP: CV leads to a higher expected presence of minority voters among those going to the polls. The results hold whether the minority is half the size of the majority or barely smaller; whether the number of open seats is just enough for CV to differentiate itself from MP ($K = 2$) or is higher ($K = 4$); whether the electorate is small or large, unless the difference in size of the two parties becomes negligible.[20] Second, the expected fraction of seats won by the minority is consistently higher under CV. The effect is most striking when the minority is relatively small ($M = 2m$), and its expected share of seats never rises above 14% under MP, less than half its share of the electorate, as opposed to being consistently close to 40% under CV.

In all cases, the minority party sets $g < K$ under CV, and thus exploits the possibility to cumulate votes. When $G = K$, the minority's cumulation of votes results in a higher probability of affecting the outcome, incentivizing turnout; when $G < K$, the difference in turnout probabilities is reduced, but the share of minority victories is boosted by the seats left uncontested by the majority.

The simulations focus on small size electorates because their purpose is to generate the hypotheses we test in the experiment. But how would the voting rules compare when the electorates are large? For MP, given its correspondence to single winner plurality systems, the theoretical predictions are known: turnout rates fall with the increase in population, but less so for the minority, whose probability of success increases with population size (for given population share) while remaining below 50 percent ([73]; [79]). There is no corresponding theoretical analysis of turnout in large populations under CV. However, the regularities we see in our simulations match the theoretical results found by Herrera et al. for proportional representation: when the relative difference in size between the two groups persists in large electorates, the difference in turnout between the minority and the majority is consistently larger under proportional representation than under plurality. In our simulations, we observe the same results under CV when $M = 2m$, for larger values of M.[21] As mentioned earlier, the logic behind the turnout decision is different under proportional

[20] If $M = m + 1$ and M is large, turnout equalizes for the two parties under both MP and CV.

[21] We have also run additional simulations with $M = 30$ and $M = 40$, confirming the qualitative results.

representation and CV, but in both cases the comparison to MP reflects the smaller impact of a large electorate on the minority decision, because the marginal impact of an additional vote is larger in a smaller group in the case of proportional representation, or because of the positive impact of cumulated votes on pivotality in the case of CV.[22]

Finally, we can compare the results to minority victories in the absence of voting costs, and thus of turnout effects. Under MP and costless voting, as we know, the minority never wins any seat, as opposed to the small but positive share predicted with costly voting. Under CV and costless voting, the expected share of minority victories is $1/2$ if $m = M - 1 > K/2$, and either $1/3$ (if $K = 2$) or $1/4$ (if $K = 4$) if $m = M/2$, as opposed to being close to 40% in all such cases if voting is costly. Changing the voting system from MP to CV always helps the minority, but the dramatic effect expected if voting is costless is mitigated when voting is costly and turnout is not universal. When voting is costly, the minority achieves substantive representation under CV but is always expected to maintain its minority status in the allocation of seats.

3.4 The Experiment

The experiment reproduces exactly the theoretical model. Our main focus is the impact of the voting rule on turnout, and especially on differential minority-majority turnout, and on the fraction of minority victories. To evaluate the robustness of the results and to test the power of the theoretical framework we implemented four different parametrizations: while we kept $M = 4$ throughout the experiment, we varied m between 2 and 3; for each m, we set $K = 2$ and $K = 4$. In all treatments, voting costs were drawn independently across participants from a uniform distribution with support $[0, 100]$, and V, the value of controlling all positions, was set at 400.

The number of candidates fielded by each party under CV was set at the theoretical equilibrium value for each parameterization and we constrained voters who turned out to spread their votes equally over their party candidates. With our parameters, not only is equal spreading of votes

[22] [89] study pivot probabilities in large electorates when voters can choose to cast more than a single vote. The problem analyzed–simultaneous referenda over multiple binary decisions–is different, but the effect of cumulation on pivot probabilities seems likely to generalize.

M, m	K	Rule	G, g	τ_{m}	τ_{M}	$(\tau_{\text{m}}$ - $\tau_{\text{M}})$	Expected Share Minority Seats
4, 2	2	MP	2, 2	0.54	0.66	-0.12	0.11
4, 2	2	CV	2, 1	0.67	0.30	0.37	0.42
4, 2	4	MP	4, 4	0.54	0.66	-0.12	0.11
4, 2	4	CV	3, 2	0.42	0.30	0.12	0.38
4, 3	2	MP	2, 2	0.56	0.79	-0.23	0.12
4, 3	2	CV	2, 1	0.49	0.27	0.22	0.42
4, 3	4	MP	4, 4	0.56	0.79	-0.23	0.12
4, 3	4	CV	3, 3	0.53	0.54	-0.01	0.42

Table 3.1: *Experimental Predictions.* F uniform over $[0, 100]$; $V = 400$.

an equilibrium response, but in the absence of distinguishing features among candidates or seats, a voter's unequal distribution of votes among the party's candidates could only reflect noise.[23] Participants acted as eligible voters: at each round, each drew an independent voting cost and decided whether or not to vote. The design thus mimics the numerical simulations, with $M = 4$. Denoting by τ_p the turnout rate of voters from party p, we reproduce in Table 3.1 the theoretical predictions for the experimental parametrizations.

The table replicates results from Figure 3.1, reporting precise numerical values. It dictates the hypotheses to test. Our main focus is on comparative predictions on the effect of a change in voting rule, from MP to CV, on the minority versus the majority. We will discuss three hypotheses in particular. First, the differential of minority-majority turnout rates $(\tau_{\text{m}} - \tau_{\text{M}})$ is strictly higher under CV than under MP. Second, a related but tighter hypothesis states that such differential is strictly negative under MP in all parametrizations; it is strictly positive under CV in three of the four parametrizations, and barely negative in the fourth. Third, in all parametrizations the expected share of seats won by the minority is higher under CV than under MP. CV benefits the minority through the difference in turnout and, when $K = 4$, because the majority should not, and in our experiment does not, contest all seats.

We conducted the experiment between August and October 2020, with participants recruited

[23]Note that there is no communication among voters.

using the Columbia Experimental Laboratory for the Social Sciences (CELSS)' ORSEE website[24]. Most subjects were undergraduate students at Columbia University or Barnard College. All sessions were online due to the COVID-19 pandemic: participants received instructions and communicated with experimenters using the Zoom videoconferencing software, and accessed the experiment interface on their personal computer's web browser. The experiment was programmed in z-Tree ([90]) and run virtually using z-Tree unleashed ([91]). Each experimental session lasted about 90 minutes with average earnings of $23. With the exception of a more visual style for the instructions, the experiment developed very similarly to in-person experiments in the lab.[25]

During each session, party sizes were kept fixed, and participants played 15 consecutive rounds of each of four treatments, CV and MP for each of $K = 2$ and $K = 4$. Having multiple treatments within a session has two main purposes: it provides some control over idiosyncratic individual behavior and, equally important in this type of experiment, keeps the subjects engaged in what is otherwise a monotonous series of decisions. We controlled for the exact sequence of the treatments by varying their order: for given m, either 2 or 3, we ran two experimental sessions for each of four orders of treatments. Thus eight sessions were conducted with $m = 2$ (12 subjects per session), and eight with $m = 3$ (14 subjects per session), for a total of 208 experimental subjects. Table 3.2 reproduces the experimental design.

Party affiliations were kept constant within each treatment to facilitate learning but were assigned randomly across treatments. In each round, two groups were formed randomly, each composed of m minority and M majority members. At the end of the round, an outcome screen reported the party affiliations of the K winning candidates and the number of members of each party who had voted. Each participant's final earnings corresponded to the sum of their earnings from one randomly drawn round from each treatment (in addition to the $5 show-up fee).

[24][46].

[25]The online appendix contains a reproduction of the instructions.

Sessions	m	Subjs	Rounds	K	Rule	Order	Sessions	m	Subjs
1,9	2	12	15	2	CV	1	5,13	3	14
			15	2	MP				
			15	4	MP				
			15	4	CV				
2,10	2	12	15	2	MP	2	6,14	3	14
			15	2	CV				
			15	4	CV				
			15	4	MP				
3,11	2	12	15	4	MP	3	7,15	3	14
			15	4	CV				
			15	2	CV				
			15	2	MP				
4,12	2	12	15	4	CV	4	8,16	3	14
			15	4	MP				
			15	2	MP				
			15	2	CV				

Table 3.2: *Experimental Design.*

3.5 Experimental Results

We begin by summarizing experimental results in Table 3.3, in the same format used for Table 3.1, for ease of comparison to the theoretical predictions.[26] We then discuss in more detail the experimental results on aggregate turnout rates–the most novel contribution of this study–and on the share of seats won by the minority–the core outcome variable. We conclude by analyzing the source of these outcomes, i.e. individual turnout decisions.

Comparing the results to Table 3.1, some observations are immediate. First, with the exception of the minority when $m = 2$ and $K = 2$, aggregate turnout in the lab under MP is close to the theory in all parametrizations and for both parties; under CV, on the other hand, turnout is consistently and substantially higher than predicted in both parties, with larger disparity for the majority. Second, in line with the theory, changing the voting rule from MP to CV leads to an increase in differential minority-majority turnout: although the quantitative effect is smaller than predicted, $(\tau_{\mathsf{m}} - \tau_{\mathsf{M}})$ is

[26]The table is descriptive and standard errors reported here are not corrected for possible correlations.

M, m	K	Rule	G, g	τ_{m}	τ_{M}	(τ_{m} - τ_{M})	Average Share Minority Seats
4, 2	2	MP	2, 2	0.40	0.64	-0.23	0.11
				(0.022)	(0.015)	(0.029)	(0.018)
4, 2	2	CV	2, 1	0.64	0.64	0.00	0.26
				(0.022)	(0.015)	(0.029)	(0.016)
4, 2	4	MP	4, 4	0.59	0.62	-0.02	0.16
				(0.022)	(0.016)	(0.027)	(0.019)
4, 2	4	CV	3, 2	0.72	0.66	0.06	0.35
				(0.020)	(0.015)	(0.027)	(0.008)
4, 3	2	MP	2, 2	0.57	0.69	-0.12	0.21
				(0.018)	(0.015)	(0.024)	(0.023)
4, 3	2	CV	2, 1	0.65	0.57	0.07	0.40
				(0.018)	(0.016)	(0.024)	(0.013)
4, 3	4	MP	4, 4	0.56	0.75	-0.20	0.13
				(0.018)	(0.014)	(0.022)	(0.019)
4, 3	4	CV	3, 3	0.62	0.74	-0.12	0.35
				(0.018)	(0.014)	(0.024)	(0.012)

Table 3.3: *Experimental Results: Summary Statistics*. Standard errors are in parentheses.

always higher under CV than under MP. In fact, as the theory predicts, the differential is negative in all MP parametrizations and positive, if weakly so, in three of the four CV parametrizations. Third, again as predicted, the share of seats captured by the minority is consistently higher under CV. We examine these results in more detail in what follows, and find them robust. They are a good summary of the experiment's main lessons.

3.5.1 Turnout

Figure 3.2 reports aggregate turnout rates for minority and majority voters in the different treatments. The darker blue columns refer to the experimental data; the lighter grey columns to the theoretical predictions, calculated from the realized voting cost draws. To account for the presence of multiple decisions by the same participant, the 95% confidence intervals are calculated from 10,000 Monte Carlo simulations that allow for correlation in turnout decisions at the individual level.[27]

For all experimental values of K and m, minority turnout is higher under CV than under MP. The effect is particularly strong for $m = 2$, but remains positive, if more muted, for $m = 3$. In the case of the majority, turnout is effectively unchanged under CV or MP, with the exception of $m = 3$ and $K = 2$, where we see a decline under CV. In this latter parametrization, the minority fields one candidate, and the majority two; the majority is certain of one victory but, with a relatively large minority concentrating all votes on a single candidate, the chances of a second majority victory are low, discouraging turnout.

In fact, the observed decline in majority turnout when $m = 3$ and $K = 2$ is much smaller than the theory predicts. The disparity is less pronounced in the other parametrizations, but, with the exception of the minority when $m = 2$ and $K = 2$, turnout under CV is consistently higher than theory predicts for both parties. In particular, the grey columns in the figure, reporting the

[27] We populate the simulations with a subject's full set of 15 choices for the relevant treatment. Each simulation corresponds to drawing, with replacement, the correct number of minority and majority subjects corresponding to the treatment, with all their decisions, and generating one number for differential turnout (the difference in the frequency of Vote decisions in each party). Over 10,000 simulations, we construct a distribution of differential turnout. The 95% CI's correspond to the boundaries of the 95% probability mass centered on the empirical ratio. Results remain substantively unchanged if we use standard automated programs for clustering standard errors.

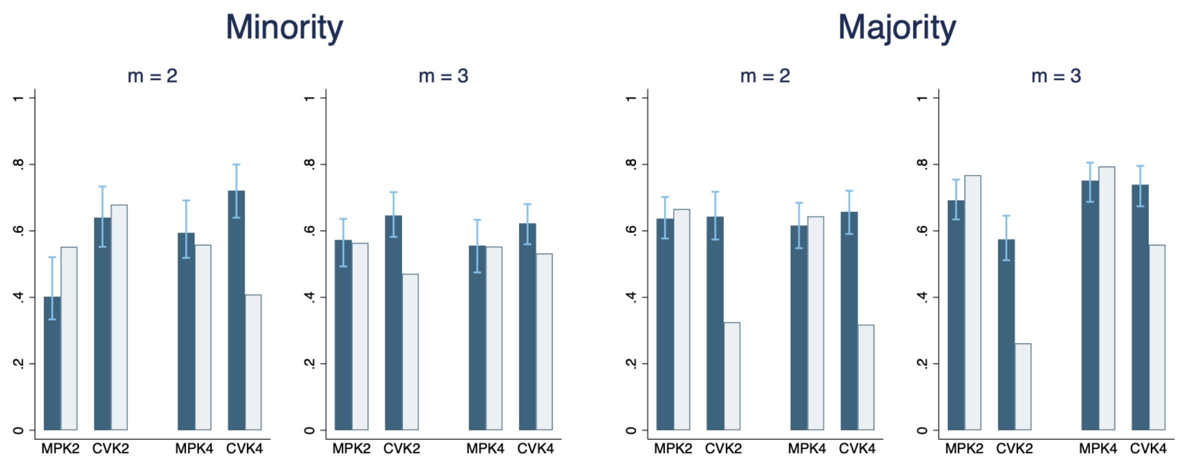


Figure 3.2: *Turnout Frequencies.* The darker blue columns correspond to the data, the lighter grey ones to the theory. The 95% CI's are calculated from 10,000 Monte Carlo simulations that allow for correlation in turnout decisions at the individual level.

theoretical predictions, show an expected decline in turnout for members of both parties in all parametrizations (again with the only exception of the minority when $m = 2$ and $K = 2$). Such generalized decline in turnout is absent from the data. The robustness of turnout under CV is the most unexpected finding of the experiment, and we return to it in Section 3.5.3, when we look in more detail at individual turnout decisions.

Given our focus on the potential role of CV in supporting the minority, the core theoretical prediction on turnout is the increase CV is expected to induce in *differential* minority to majority turnout rates ($\tau_{\mathrm{m}} - \tau_{\mathrm{M}}$). Figure 3.3 reports the experimental results (the darker diamonds) and the theoretical predictions (the lighter diamonds). In all cases, the differential minority to majority turnout is higher under CV than under MP.

The absolute magnitudes of the experimental effects are muted, relative to the theory, particularly in the $K = 2$ treatments, reflecting the higher than expected majority turnout. Comparative magnitudes, however, are roughly in line with predictions—more pronounced in $K = 2$ treatments, and more muted when $K = 4$. The theory also predicts that the minority's turnout rate should be lower than the majority's in all MP treatments, higher in three of the four CV treatments, and barely lower in the fourth. This too is observed in the experimental results.

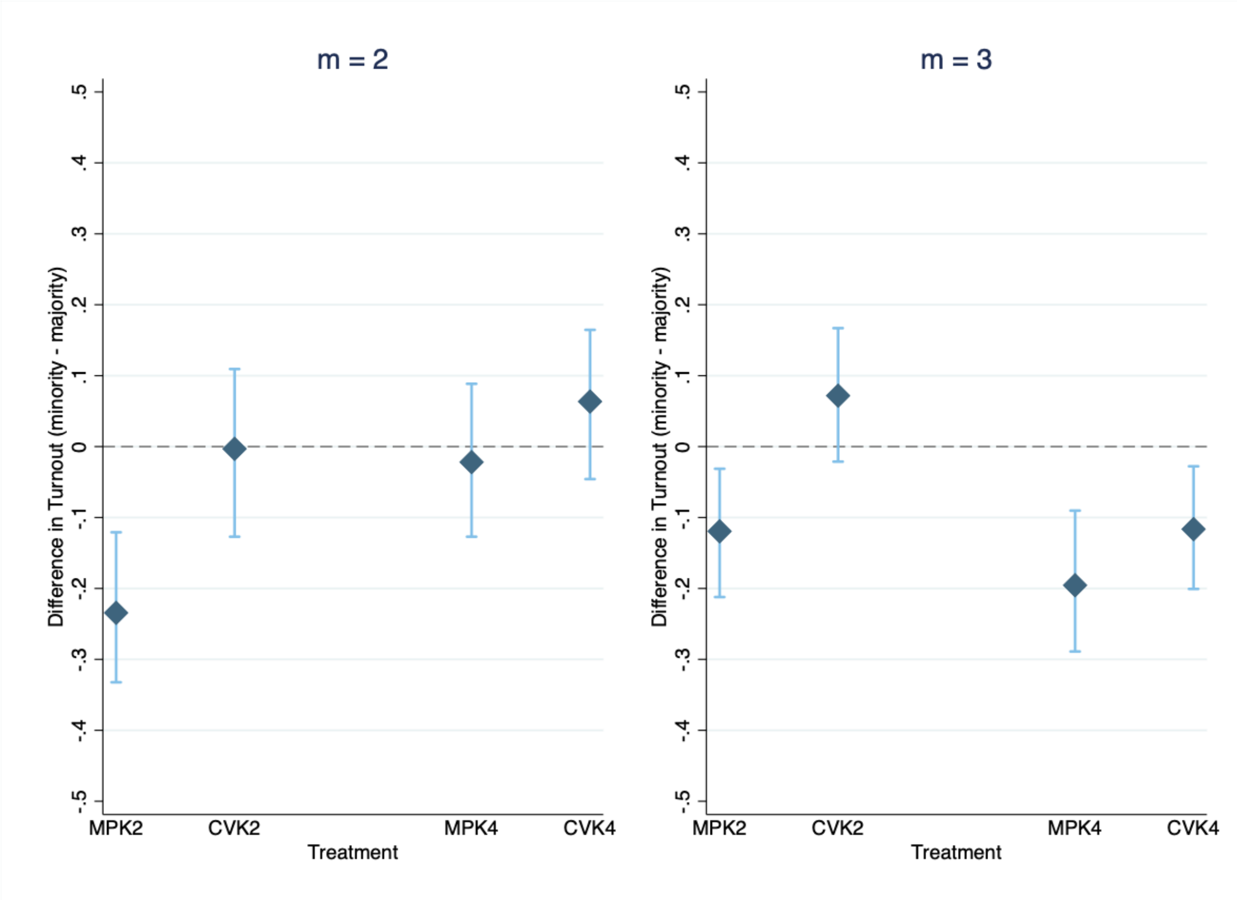


Figure 3.3: *Differential Turnout*: $\tau_{\mathrm{m}} - \tau_{\mathrm{M}}$. The 95% CI's are calculated from 10,000 Monte Carlo simulations that allow for correlation in turnout decisions at the individual level.

We construct a non-parametric test of the significance of these results by comparing them to the corresponding simulations under the null hypothesis of no difference in behavior between minority and majority voters. For each treatment t—i.e., for each parametrization and voting rule—we call $n_1(t)$ the size of the minority sample, and $n_2(t)$ the size of the majority sample. We then combine minority and majority subjects in a single sample and construct two random groups, 1 and 2, labeled $g_1(t)$ and $g_2(t)$, by drawing subjects with replacement from the joint sample and assigning $n_1(t)$ random draws to group 1 and $n_2(t)$ to group 2.[28] We treat the samples in the two groups as if they corresponded to the minority and to the majority, and calculate differential turnout $(\tau_1 - \tau_2)$. We repeat the procedure 10,000 times and obtain a distribution of differential turnout, under the

[28]To allow for correlation of individual decisions, each subject is drawn with all 15 rounds of turnout decisions and voting costs.

hypothesis of no systematic difference in turnout between minority and majority subjects. The relevant p-value is the probability mass of the distribution at $(\tau_1 - \tau_2) \geq (\tau_m - \tau_M)$, where the latter is the difference in turnout rates observed in the data. We find: $p = 0.003$ ($m = 2, K = 2$; and $m = 3, K = 2$), $p = 0.137$ ($m = 2, K = 4$), $p = 0.135$ ($m = 3, K = 4$). The test confirms the visual evidence of the figure: the impact of CV on differential minority-majority turnout is clear in the $K = 2$ parametrizations; it is still present but smaller and thus less sharply identified when $K = 4$.

Further analysis supports these conclusions. In Table 3.4, for given parametrization, we report the estimation of a linear probability model where the subjects' turnout decisions are regressed on a dummy variable for minority status, a dummy variable for CV, and the interaction of the two dummies, controlling for voting costs, for the relative order of CV and MP, and for the round number.

The table tells us that, in all parametrizations, belonging to the minority decreases turnout under MP, relative to the majority voters, while shifting to CV has no impact on majority voters' turnout, with the exception of the case $m = 3$, $K = 2$, when the majority turnout declines. At the center of our predictions, the positive parameter of the interaction term tells us that shifting to CV has a larger positive effect (or a smaller negative effect) on the turnout of minority voters relative to majority voters, The effect is precisely estimated when $K = 2$, although not in the more complex treatments with $K = 4$. As expected, higher voting costs decrease turnout, but whether subjects see CV before or after MP has no noticeable effect on the decision to vote, although as a session proceeds, turnout slightly decreases.

These qualitative results are robust. As we report in the online Appendix, they remain unchanged if we estimate the model using only data where the relevant treatment was the one the subjects saw first—i.e., where we exploit the changing order of treatments across sessions to create an in-between subjects design. Results also remain unchanged if we estimate a probit model, rather than a linear probability model, with or without random effects.[29]

[29] Interacting the order of treatments with CV (i.e., adding a variable CV × CVfirst) has no effect. Changing the level at which standard errors are clustered affects statistical significance but does not change the main message. Focusing on the interaction term CV × Minority, and thus on the effect of CV on differential turnout, clustering at the group level leads to a decline in all estimated standard errors; clustering at the session level increases estimated standard

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
Minority	-0.229***	-0.042	-0.144***	-0.202***
	(0.052)	(0.052)	(0.045)	(0.052)
CV	0.008	0.029	-0.130***	-0.027
	(0.035)	(0.035)	(0.038)	(0.030)
CV × Minority	0.227***	0.118	0.218***	0.091
	(0.078)	(0.072)	(0.069)	(0.061)
Voting Cost	-0.007***	-0.008***	-0.009***	-0.008***
	(0.001)	(0.001)	(0.000)	(0.001)
CVfirst	-0.057	-0.029	0.029	-0.006
	(0.047)	(0.047)	(0.040)	(0.040)
Round	-0.008***	-0.003	-0.008***	-0.005***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	1.095***	1.048***	1.201***	1.189***
	(0.039)	(0.047)	(0.033)	(0.037)
Observations	2880	2880	3360	3360
R^2	0.239	0.228	0.296	0.268

Standard errors in parentheses. 96 clusters in columns 1,2; 112 in columns 3,4.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: *Individual turnout decisions.* The default sample is majority subjects under MP. Standard errors are clustered at the individual level.

The regression results mirror closely what we see in the figures. Lower minority turnout under MP is less precisely estimated when $m = 2$ and $K = 4$; the decline in majority turnout under CV is only detectable for the case $m = 3$, $K = 2$, and the positive effect of CV on differential minority-majority turnout is strongly significant for the $K = 2$ treatments, but not when $K = 4$. These are the regularities we also see in the figures.

We can use the simulations reported earlier, estimating differential turnout under the null hypothesis of no difference in behavior across minority and majority subjects, to construct a placebo test. The simulations yield turnout rates for group 1 and group 2, the two groups generated ran-

errors for the $m = 2$ parametrizations and reduces them when $m = 3$. Clustering at the individual level, as in Table 3.4, is thus more conservative than the former and less conservative than the latter. See the online Appendix for all alternative results and more discussion.

domly from our data, for each treatment, by drawing subjects without distinguishing by party. For each simulation, we replicate the regression in Table 3.4; over 10,000 simulations, we generate a distribution of the parameter of the interaction term, CV $\times$ group 1, under the null of no difference in the population. Figure 3.4 reports such distribution for each parametrization, as well as the 95% confidence interval around the distribution mean, and, with a thicker black line, the parameter estimated in the original regression.

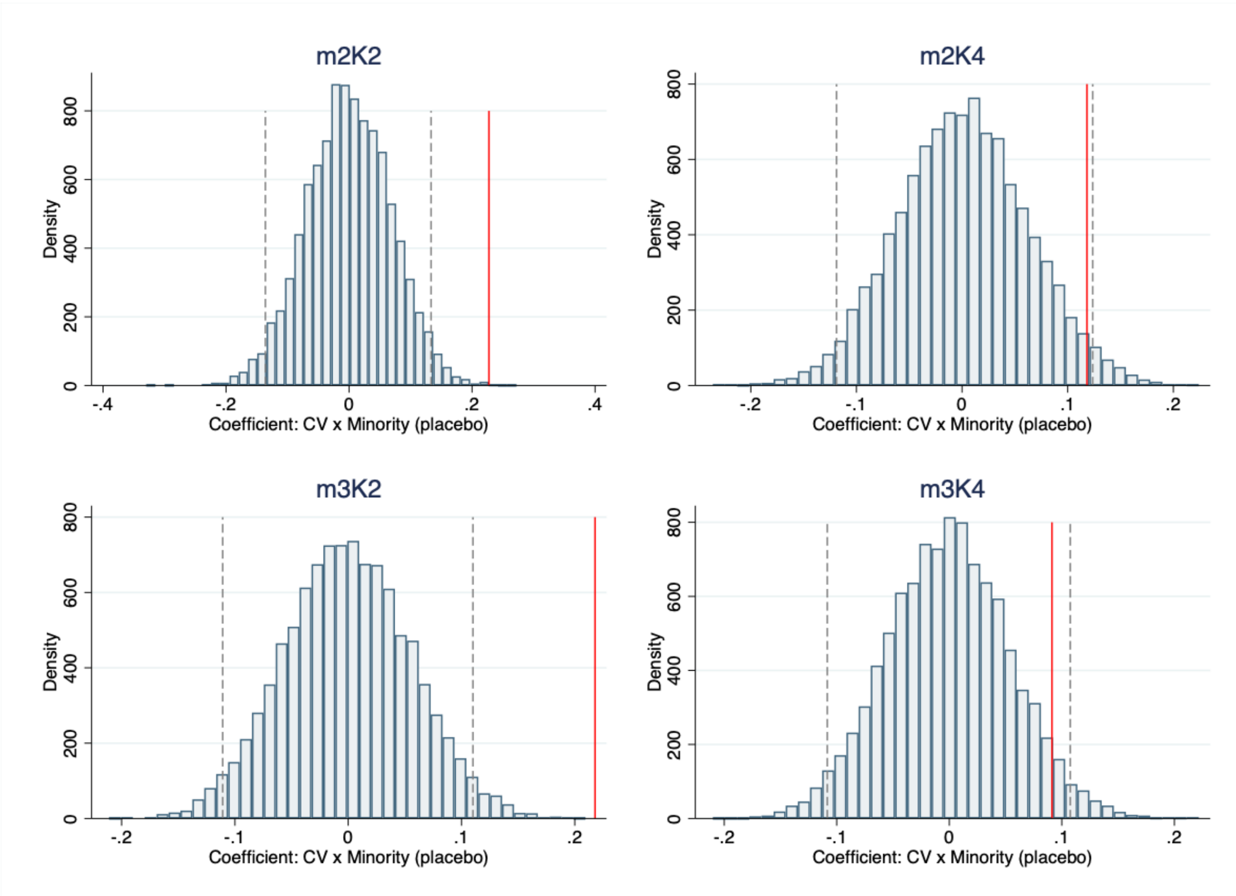


Figure 3.4: *The effect of CV on differential minority-majority turnout rates. A placebo test*. Results of the original regressions (the thicker black lines) versus 10,000 replications with random group formation. The dotted lines correspond to the 95% CI around the distribution mean.

Under the null of no systematic differences in turnout decisions between minority and majority voters, the probability of estimating a differential effect of CV on the minority equal or larger to the estimate in our original regression is effectively zero in both $K = 2$ treatments ($p < 0.001$); such probability is slightly larger in the $K = 4$ treatments but still below conventional significance

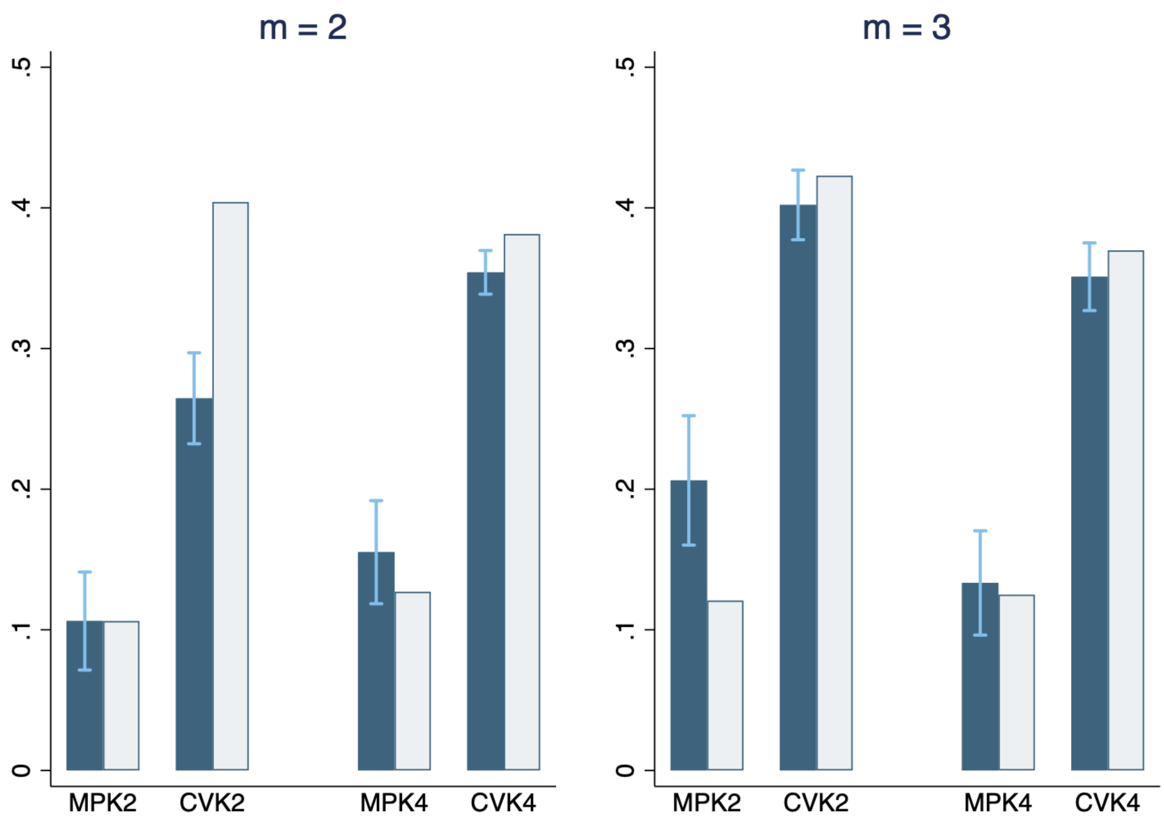


Figure 3.5: *Share of seats won by the minority.* Darker blue columns correspond to the data, lighter grey to the theory. The 95% CI's are calculated from standard errors clustered at the level of the voting group.

levels ($p = 0.031$ when $m = 2$; $p = 0.048$ when $m = 3$). The placebo test suggests strongly that minority and majority members do in fact respond differently to the shift in voting rule from MP to CV. As a result, differential minority-majority turnout is higher under CV than under MP, an effect we cannot attribute to randomness.

3.5.2 Minority victories

Did CV help the minority secure more seats? Figure 3.5 shows that the answer is positive.

For every parametrization, CV increases the share of seats won by the minority, and does so very significantly. There are some disparities relative to the theory: the minority fares better than expected under MP in the $m = 3$, $K = 2$ treatment, and less well than expected under CV in the

$m = 2$, $K = 2$ treatment. On the whole, however, the results are in line with predictions: not only do minority victories increase under CV in all cases, but the magnitude of the change is large: in all parametrizations, the share of seats won by the minority doubles or more than doubles when shifting from MP to CV.[30]

The message of the figure is confirmed by the statistical analysis. Table 3.5 reports the results of regressing, for each parametrization, the share of seats won by the minority on a dummy variable for CV, controlling for the relative order of CV and MP, for an interaction term between the order and CV (CV×CVfirst), and for the round number. In all parametrizations, CV consistently and substantially increases the share of minority victories. The order of treatments per se is not significant, although in the case of the $m = 2$ and $K = 2$ parametrization, the most transparent, the effect of CV on increasing minority victories is stronger in sessions where CV was run after MP. Regressions run on data restricted to the first treatment in each session, however, confirm that the positive and significant effect of CV on minority victories is very robust: as we show in the online Appendix, the finding is confirmed in such restricted data set for all treatments, including when $m = 2$ and $K = 2$.[31] Under CV, the possibility to cumulate votes combines with the increase in differential minority-majority turnout to deliver larger influence to the minority.

3.5.3 Individual turn out decisions: monotonicity violations and cutpoints

Aggregate group outcomes–turnout rates and shares of seats won–are the main variables of interest. But group outcomes are rooted in individual turnout decisions. In this subsection, we briefly discuss the experimental evidence on individual behavior.

In cost of voting experiments, and more broadly in experiments where the equilibrium is in monotone cutpoint strategies, violations of monotonicity are informative not only about the accuracy of the theoretical predictions but also about the participants' understanding of the rules of the game. In our experiment, this is particularly important because a common objection to CV is

[30] Although, as predicted, such share remains under 50% in all cases.

[31] See the online Appendix, where we also document that results are unchanged when clustering standard errors at the session level.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
CV	0.241***	0.202***	0.230***	0.224***
	(0.032)	(0.031)	(0.037)	(0.028)
CVfirst	0.045	-0.003	0.069	0.059
	(0.036)	(0.037)	(0.047)	(0.038)
CV × CVfirst	-0.164***	-0.005	-0.068	-0.011
	(0.050)	(0.040)	(0.054)	(0.045)
Round	-0.001	-0.000	-0.002	0.001
	(0.003)	(0.002)	(0.003)	(0.002)
Constant	0.094**	0.161***	0.188***	0.098***
	(0.036)	(0.032)	(0.039)	(0.030)
Observations	1649	1641	1870	1861
R^2	0.094	0.165	0.082	0.131

Standard errors in parentheses.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: *Share of seats won by the minority.* Standard errors are clustered at the level of the voting group.

that its strategic complexity is a difficult obstacle for voters. Although experimental participants limit themselves to the decision to turn out or not, the fact that under CV turnout implies casting multiple votes for each candidate, and in one case fractional votes, could be confusing.

For each participant, we calculated the number of monotonicity violations, defined as the minimum number of decisions that would need to be modified for that participant's turnout to be fully monotonic in the voting cost realization: if i chooses to turn out for a cost realization $c_i = c'$ then i should turn out for all $c_i < c'$, and if i chooses to abstain for a cost realization $c_i = c''$ then i should abstain for all $c_i > c''$. In our data, monotonicity violations are not common: in all treatments more than half of participants have at most a single violation. The statement in fact undersells the evidence: over the full data set, 75 percent of participants have at most one violation. What is most important for our purposes, there is no systematic difference between the frequency of violations under MP and under CV. As we show more formally in the online Appendix, at least in the simplified structure of our experiment, the hypothesis that CV is more confusing for voters is not supported by any evidence of more random behavior.[32]

We can use the minimization of monotonicity violations as a guide to estimating individual cost cutpoints. Figure 3.6 reports, for each subject, the cutpoint that minimizes the frequency of violations.[33] Because the theory predicts different behavior depending on party affiliation, the figure reports estimated cutpoints separately for each party. In all cases, the darker blue diamonds correspond to the average of the individual cutpoints, and the lighter grey diamond to the theoretical prediction.

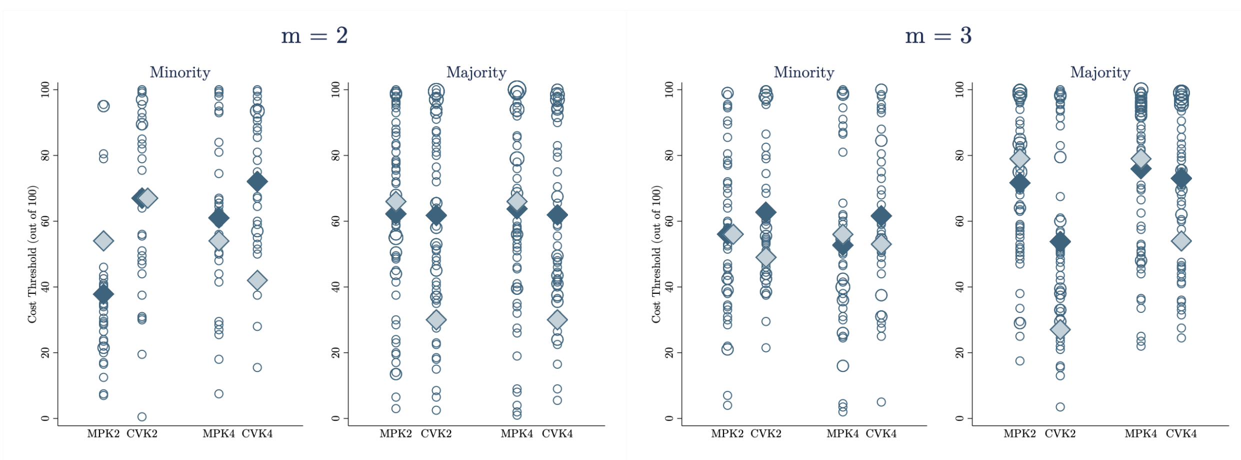


Figure 3.6: *Estimated Cost Cutpoints*. The darker blue diamonds correspond to the average of the individual cutpoints, and the lighter grey diamonds to the theoretical prediction. The size of each circle is proportional to the number of subjects it represents.

The figure shows clearly the high heterogeneity in behavior: estimated cutpoints vary across individuals, in both parties and for both voting rules. The theoretical semi-symmetric equilibrium predicts a single cutpoint for each party, a prediction clearly violated by the data. The heterogeneity we see, however, is in line with previous findings from similar experiments.[34] Comparing the

[32]We report in the online Appendix the fraction of individuals with different numbers of monotonicity violations, represented as separate CDF's for MP and CV. The two CDF's are barely distinguishable, and Kolmogoroff-Smirnov (K-S) two-sample tests, corrected for discreteness, cannot reject the hypothesis of a common population,

[33]See, for example, [92] or [73] for a similar approach. When, for a given subject, multiple cutpoints are consistent with minimizing monotonicity violations, the figure reports the mean cutpoint. In a few cases (15 subjects out of 208), the multiplicity concerns ranges of possible cutpoints; in these cases reporting the mean would muddle behavior, and we have chosen the range that is closest to equilibrium. In all cases, we have verified that alternative choices do not change the qualitative results.

[34]For example, [73]. We checked whether restricting data to the last 10 rounds of each treatment would result in a narrower range of estimates. We found no such effect: the range of estimated cutpoints is virtually identical.

dispersion of cutpoints across the two voting rules, we find very similar dispersion across voting rules and treatments, both for the minority and for the majority. The visual impression is confirmed by the standard deviations of the corresponding distributions of cutpoints.

When we aggregate individual behavior into average cutpoints, regularities emerge. Under MP, the average of the estimated individual cutpoints is remarkably close to the equilibrium cutpoint in all parametrizations, with the single exception of the minority when $m = 2$ and $K = 2$. With the same single exception, under CV, the average cutpoint is above the equilibrium cutpoint in both parties and in all parametrizations, indicating more frequent participation in voting than theory predicts. When $m = 2$ and $K = 2$, the average minority cutpoint in the data is below the equilibrium cutpoint for MP (suggesting less participation), and coincides with the equilibrium for CV.

A different way of visualizing estimated individual cutpoints conveys clearer lessons on the impact of the voting rule. Figure 3.7 reports the CDF's of the individual cutpoints, comparing the CDF's under MP (the lighter line) and under CV (the darker line), for both parties and all parametrizations.

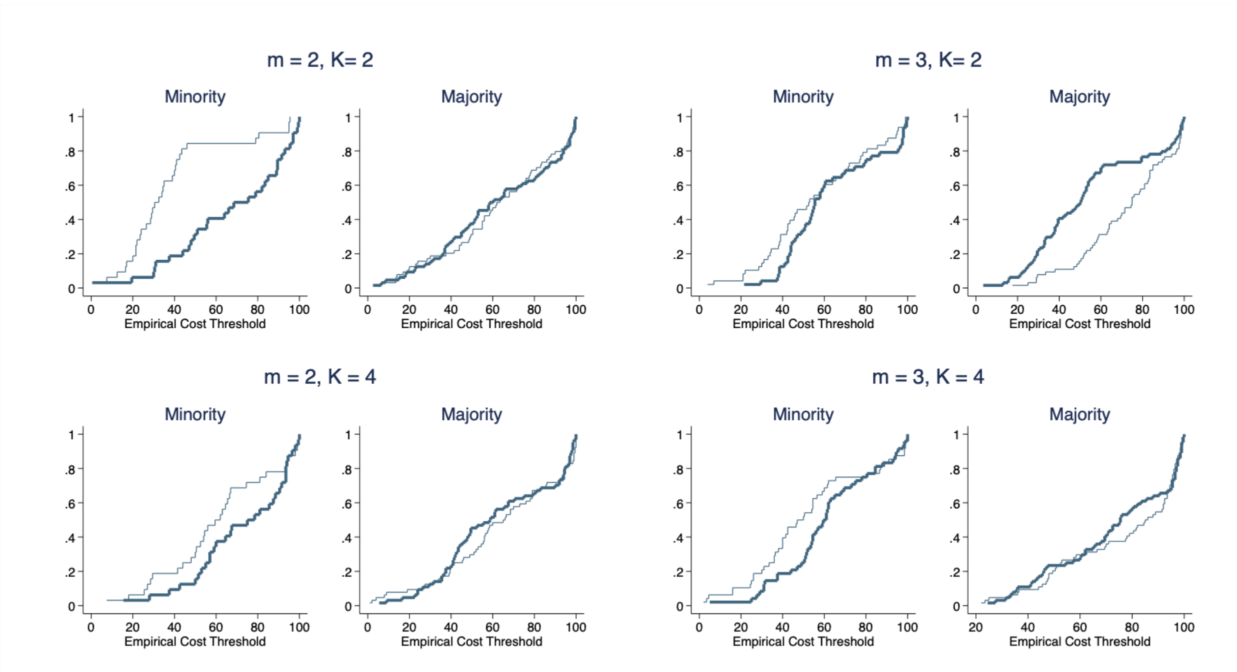


Figure 3.7: *CDF's of Cost Cutpoints.* The darker lines correspond to CV; the lighter lines to MP.

The minority's higher propensity to vote under CV operates throughout the distribution of individual cutpoints. The move to CV causes a shift rightward of the whole distribution: the minority cutpoints distribution under CV FOSD's the distribution under MP. Formal tests strongly reject the hypothesis of equal distributions for the minority when $m = 2$ and $K = 2$ (a two-sample KS test yields $p < 0.001$), with weaker evidence when $m = 3$ and $K = 4$ ($p = 0.05$) and when $m = 2$ and $K = 4$ ($p = 0.26$).

As for the majority, with $m = 2$, majority members barely modify their propensity to vote. With $m = 3$, the conclusion is similar when $K = 4$; when $K = 2$, however, we see a consistent decline in voting under CV throughout the majority's cutpoint distribution. When $m = 3$, $K = 2$, as we discussed earlier, under CV the majority's chances of winning both seats are very small and that realization depresses turnout: the distribution of majority cutpoints under MP FOSD's the distribution under CV, and the corresponding one-sided KS test yields $p < 0.001$.[35]

The analysis of individual behavior in the experiment thus delivers four main lessons. First, we do not see more random behavior under CV than under MP, as captured by the frequency of monotonicity violations. Such violations are few under both voting rules. Second, estimated individual cutpoints are heterogeneous, with comparable dispersion across parametrizations, parties, and voting rules. Third, but for a single exception (the minority, when $m = 2$, and $K = 2$), average cutpoints are close to the theoretical predictions for MP, but consistently higher than predicted under CV, especially, but not exclusively, for the majority. Fourth, plotting CDF's of cutpoints highlights participants' differential behavior under the two voting rules. The distribution of minority party subjects' cutpoints shifts to the right (i.e., towards higher turnout) under CV, relative to MP; it moves less, and when it does it is in the opposite direction, for majority party subjects.

3.6 Conclusions

Cumulative Voting (CV) is a voting system for multi-member districts that allows each voter to cumulate votes freely on a single or a subset of candidates. It has been in use since the 19th

[35] The test cannot reject the hypothesis of equal distributions of majority cutpoints in the other cases.

century in different countries, for example in the US, England, and Scotland, for the election of both political bodies and corporate boards. Because it delivers semi-proportional outcomes without imposing a proportional representation system, CV favors the representation of minorities while remaining familiar and acceptable to majoritarian democratic systems. In the US, it is one of the remedies imposed by the Courts to resolve violations of the Voting Rights Act. Adoption of CV correlates empirically with higher minorities participation and success. However, because such adoption typically follows litigation, and thus heightened engagement of minorities, it is important to complement the historical experiences with experiments. This is the purpose of the present study.

We use a traditional experimental design with costly voting to test predictions on turnout and on the electoral success of the minority under two voting systems: Multiwinner Plurality (MP), or bloc voting, where each voter can only cast one vote per candidate, and CV. Inspired by historical episodes, we assume that the coordination problem posed by CV is solved by the party leaders. The leaders choose the number of party candidates optimally, in an equilibrium in which votes are spread evenly over all party candidates. Experimental subjects are voters who face a private, individual cost of voting, and must decide whether or not to turn out.

Although we see more individual heterogeneity in behavior than expected, aggregate data under MP are well predicted by the theoretical model. We find, however, that the model significantly underpredicts turnout under CV for both parties, with a larger quantitative discrepancy for the majority. Given the good performance of the model under MP, the finding is puzzling. Why the underprediction under CV only? Without claiming to have an answer, one possible conjecture is that experimental subjects correctly perceive CV as a more competitive system, giving the minority a higher chance of winning seats. The enhanced competition then leads to higher participation, in the same spirit as overbidding in auction experiments.[36] We leave the conjecture open for future work.

The higher turnout under CV does not invalidate the model's predictions about the relative

[36] A variation of the "spite" argument invoked to explain overbidding in second price auctions ([93]), could then justify the majority's higher excess turnout.

impact of the voting system on minority and majority voters. For the cases we bring to the lab, the theory says that CV should be associated with an increase in the turnout of the minority, relative to the majority, and with an increase in the share of seats won by the minority. Both predictions are supported by our data in all the experimental parametrizations we study. In our experiment, CV works as expected, to magnify the voice of the minority.

As we write, debates over voting rights rage in Congress, in state legislatures, in the courts, in the media. Initiatives aimed at limiting access to the polls and partisan redistricting following the 2020 census increase fears of disenfranchisement. Encouraging minority turnout is a high priority. Voting rules like CV have the potential to help. Although much remains to be studied, in the lab we find that such potential is fulfilled.[37]

[37]This chapter was published in Volume 141 of *Games and Economic Behavior*, pp. 133-155, September 2023, https://doi.org/10.1016/j.geb.2023.05.012.

References

[1] S. Frenkel and K. Browning, "The Metaverse's Dark Side: Here Come Harassment and Assaults," *The New York Times*, Dec. 2021.

[2] Pew Research Center, "The State of Online Harassment," 2021.

[3] J. Andreoni and R. Petrie, "Public goods experiments without confidentiality: A glimpse into fund-raising," *Journal of Public Economics*, vol. 88, no. 7, pp. 1605–1623, Jul. 2004.

[4] M. Rege and K. Telle, "The impact of social approval and framing on cooperation in public good situations," *Journal of Public Economics*, vol. 88, no. 7, pp. 1625–1644, Jul. 2004.

[5] T. Groves and J. Ledyard, "Optimal Allocation of Public Goods: A Solution to the "Free Rider" Problem," *Econometrica*, vol. 45, no. 4, pp. 783–809, 1977, Publisher: [Wiley, Econometric Society].

[6] J. O. Ledyard, "Public goods: A survey of experimental research," in *The Handbook of Experimental Economics*, J. H. Kagel and A. E. Roth, Eds., Princeton, NJ: Princeton University Press, 1995.

[7] A. Chaudhuri, "Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature," *Experimental Economics*, vol. 14, no. 1, pp. 47–83, Mar. 2011.

[8] R. M. Isaac, J. M. Walker, and S. H. Thomas, "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations," *Public Choice*, vol. 43, no. 2, pp. 113–149, 1984, Publisher: Springer.

[9] E. Fehr and K. M. Schmidt, "A Theory of Fairness, Competition, and Cooperation," *The Quarterly Journal of Economics*, vol. 114, no. 3, pp. 817–868, 1999, Publisher: Oxford University Press.

[10] J. Andreoni, "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence," *Journal of Political Economy*, vol. 97, no. 6, pp. 1447–1458, 1989, Publisher: University of Chicago Press.

[11] J. Andreoni, "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving," *The Economic Journal*, vol. 100, no. 401, pp. 464–477, 1990, Publisher: [Royal Economic Society, Wiley].

[12] S. Gächter and E. Fehr, "Collective action as a social exchange," *Journal of Economic Behavior & Organization*, vol. 39, no. 4, pp. 341–369, Jul. 1999.

[13] A. Samek and R. M. Sheremeta, “Recognizing contributors: An experiment on public goods,” *Experimental Economics*, vol. 17, no. 4, pp. 673–690, Dec. 2014.

[14] A. Falk, “Facing yourself – A note on self-image,” *Journal of Economic Behavior & Organization*, vol. 186, pp. 724–734, Jun. 2021.

[15] R. Bénabou and J. Tirole, “Incentives and Prosocial Behavior,” *American Economic Review*, vol. 96, no. 5, pp. 1652–1678, Dec. 2006.

[16] E. Ostrom, J. Walker, and R. Gardner, “Covenants With and Without a Sword: Self-Governance is Possible,” *The American Political Science Review*, vol. 86, no. 2, pp. 404–417, 1992, Publisher: [American Political Science Association, Cambridge University Press].

[17] E. Fehr and S. Gächter, “Cooperation and Punishment in Public Goods Experiments,” *The American Economic Review*, vol. 90, no. 4, pp. 980–994, 2000, Publisher: American Economic Association.

[18] E. Fehr and S. Gächter, “Altruistic punishment in humans,” *Nature*, vol. 415, no. 6868, pp. 137–140, Jan. 2002, Number: 6868 Publisher: Nature Publishing Group.

[19] L. Denant-Boemont, D. Masclet, and C. N. Noussair, “Punishment, counterpunishment and sanction enforcement in a social dilemma experiment,” *Economic Theory*, vol. 33, no. 1, pp. 145–167, 2007, Publisher: Springer.

[20] N. Nikiforakis, “Feedback, punishment and cooperation in public good experiments,” *Games and Economic Behavior*, vol. 68, no. 2, pp. 689–702, Mar. 2010.

[21] M. Arellano and S. Bond, “Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations,” *The Review of Economic Studies*, vol. 58, no. 2, pp. 277–297, Apr. 1991.

[22] R. T. A. Croson, “Theories of Commitment, Altruism and Reciprocity: Evidence from Linear Public Goods Games,” *Economic Inquiry*, vol. 45, no. 2, pp. 199–216, 2007.

[23] J. Shang and R. Croson, “A Field Experiment in Charitable Contribution: The Impact of Social Information on the Voluntary Provision of Public Goods,” *The Economic Journal*, vol. 119, no. 540, pp. 1422–1439, Oct. 2009.

[24] S. Meier, “A survey of economic theories and field evidence on pro-social behavior.,” in *Economics and psychology: A promising new cross-disciplinary field.* Ser. CESifo seminar series. Cambridge, MA, US: MIT Press, 2007, pp. 51–87, ISBN: 0-262-06263-1 (Hardcover); 978-0-262-06263-3 (Hardcover).

[25] M. Dufwenberg and A. Muren, “Generosity, anonymity, gender,” *Journal of Economic Behavior & Organization*, vol. 61, no. 1, pp. 42–49, Sep. 2006.

[26] G. Charness and U. Gneezy, “What’s in a name? Anonymity and social distance in dictator and ultimatum games,” *Journal of Economic Behavior & Organization*, vol. 68, no. 1, pp. 29–35, Oct. 2008.

[27] D. Ariely, A. Bracha, and S. Meier, “Doing Good or Doing Well? Image Motivation and Monetary Incentives in Behaving Prosocially,” *The American Economic Review*, vol. 99, no. 1, pp. 544–555, 2009, Publisher: American Economic Association.

[28] A. R. Soetevent, “Anonymity in giving in a natural context—a field experiment in 30 churches,” *Journal of Public Economics*, vol. 89, no. 11, pp. 2301–2323, Dec. 2005.

[29] D. Karlan and M. A. McConnell, “Hey look at me: The effect of giving circles on giving,” *Journal of Economic Behavior & Organization*, vol. 106, pp. 402–412, Oct. 2014.

[30] J. Sell and R. K. Wilson, “Levels of Information and Contributions to Public Goods*,” *Social Forces*, vol. 70, no. 1, pp. 107–124, Sep. 1991.

[31] J. Weimann, “Individual behaviour in a free riding experiment,” *Journal of Public Economics*, vol. 54, no. 2, pp. 185–200, Jun. 1994.

[32] R. K. Wilson and J. Sell, “"Liar, Liar...": Cheap Talk and Reputation in Repeated Public Goods Settings,” *The Journal of Conflict Resolution*, vol. 41, no. 5, pp. 695–717, 1997, Publisher: Sage Publications, Inc.

[33] R. Croson, “Feedback in Voluntary Contribution Mechanisms: An Experiment in Team Production,” *Research in Experimental Economics*, vol. 8, pp. 85–97, Jan. 2001.

[34] M. Jones and M. McKee, “Feedback Information and Contributions to Not-for-Profit Enterprises: Experimental Investigations and Implications for Large-Scale Fund-Raising,” *Public Finance Review*, vol. 32, no. 5, pp. 512–527, Sep. 2004, Publisher: SAGE Publications Inc.

[35] S. Lim and B. Reeves, “Being in the Game: Effects of Avatar Choice and Point of View on Psychophysiological Responses During Play,” *Media Psychology*, vol. 12, no. 4, pp. 348–370, Nov. 2009, Publisher: Routledge https://doi.org/10.1080/15213260903287242.

[36] N. Yee and J. Bailenson, “The Proteus Effect: The Effect of Transformed Self-Representation on Behavior,” *Human Communication Research*, vol. 33, no. 3, pp. 271–290, Jul. 2007.

[37] V. Groom, J. N. Bailenson, and C. Nass, “The influence of racial embodiment on racial bias in immersive virtual environments,” *Social Influence*, vol. 4, no. 3, pp. 231–248, Jul. 2009, Publisher: Routledge https://doi.org/10.1080/15534510802643750.

[38] G. Bente, S. Rüggenberg, N. C. Krämer, and F. Eschenburg, “Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations,” *Human Communication Research*, vol. 34, no. 2, pp. 287–318, 2008.

[39] M. Fiedler and E. Haruvy, "The lab versus the virtual lab and virtual field—An experimental investigation of trust games with communication," *Journal of Economic Behavior & Organization*, vol. 72, no. 2, pp. 716–724, Nov. 2009.

[40] G. Bente, T. Dratsch, K. Kaspar, T. Häßler, O. Bungard, and A. Al-Issa, "Cultures of Trust: Effects of Avatar Faces and Reputation Scores on German and Arab Players in an Online Trust-Game," *PLOS ONE*, vol. 9, no. 6, e98297, Jun. 2014, Publisher: Public Library of Science.

[41] A. Vasalou and A. N. Joinson, "Me, myself and I: The role of interactional context on self-presentation through avatars," *Computers in Human Behavior*, Including the Special Issue: State of the Art Research into Cognitive Load Theory, vol. 25, no. 2, pp. 510–520, Mar. 2009.

[42] R. M. Martey and M. Consalvo, "Performing the Looking-Glass Self: Avatar Appearance and Group Identity in Second Life," *Popular Communication*, vol. 9, no. 3, pp. 165–180, Jul. 2011, Publisher: Routledge https://doi.org/10.1080/15405702.2011.583830.

[43] G. Charness, R. Cobo-Reyes, S. Meraglia, and n. Sánchez, "Anticipated Discrimination, Choices, and Performance: Experimental Evidence," *European Economic Review*, vol. 127, p. 103 473, Aug. 2020.

[44] D. Abraham, B. Greiner, and M. Stephanides, "On the Internet you can be anyone: An experiment on strategic avatar choice in online marketplaces," *Journal of Economic Behavior & Organization*, vol. 206, pp. 251–261, Feb. 2023.

[45] V. Lugovskyy, D. Puzzello, A. Sorensen, J. Walker, and A. Williams, "An experimental study of finitely and infinitely repeated linear public goods games," *Games and Economic Behavior*, vol. 102, pp. 286–302, Mar. 2017.

[46] B. Greiner, "Subject pool recruitment procedures: Organizing experiments with ORSEE," *Journal of the Economic Science Association*, vol. 1, no. 1, pp. 114–125, Jul. 2015.

[47] D. L. Chen, M. Schonger, and C. Wickens, "oTree—An open-source platform for laboratory, online, and field experiments," *Journal of Behavioral and Experimental Finance*, vol. 9, pp. 88–97, Mar. 2016.

[48] R. Bénabou and J. Tirole, "Identity, Morals, and Taboos: Beliefs as Assets," *The Quarterly Journal of Economics*, vol. 126, no. 2, pp. 805–855, 2011, Publisher: Oxford University Press.

[49] P. B. Seabright, "Continuous Preferences and Discontinuous Choices: How Altruists Respond to Incentives," *The B.E. Journal of Theoretical Economics*, vol. 9, no. 1, Apr. 2009, Publisher: De Gruyter.

[50] D. S. Hamermesh and J. E. Biddle, "Beauty and the Labor Market," *The American Economic Review*, vol. 84, no. 5, pp. 1174–1194, 1994, Publisher: American Economic Association.

[51] T. S. Rosenblat, "The Beauty Premium: Physical Attractiveness and Gender in Dictator Games," *Negotiation Journal*, vol. 24, no. 4, pp. 465–481, 2008.

[52] C. Kleinhans and N. Nicholls, "Beautiful inside and out? The role of physical attractiveness in predicting altruistic behaviour," *International Social Science Journal*, vol. 73, no. 248, pp. 613–626, 2023, https://onlinelibrary.wiley.com/doi/pdf/10.1111/issj.12422.

[53] J. Andreoni and R. Petrie, "Beauty, gender and stereotypes: Evidence from laboratory experiments," *Journal of Economic Psychology*, vol. 29, no. 1, pp. 73–93, Feb. 2008.

[54] P. Mussel, A. S. Göritz, and J. Hewig, "The value of a smile: Facial expression affects ultimatum-game responses," *Judgment and Decision Making*, vol. 8, no. 3, pp. 381–385, May 2013.

[55] M. Tonin and M. Vlassopoulos, "Experimental evidence of self-image concerns as motivation for giving," *Journal of Economic Behavior & Organization*, vol. 90, pp. 19–27, Jun. 2013.

[56] R. K. Wilson and C. C. Eckel, "Judging a Book by its Cover: Beauty and Expectations in the Trust Game," *Political Research Quarterly*, vol. 59, no. 2, pp. 189–202, Jun. 2006, Publisher: SAGE Publications Inc.

[57] L. Litman, J. Robinson, and T. Abberbock, "TurkPrime.com: A versatile crowdsourcing data acquisition platform for the behavioral sciences," *Behavior Research Methods*, vol. 49, no. 2, pp. 433–442, Apr. 2017.

[58] D. J. Hauser, A. J. Moss, C. Rosenzweig, S. N. Jaffe, J. Robinson, and L. Litman, "Evaluating CloudResearch's Approved Group as a solution for problematic data quality on MTurk," *Behavior Research Methods*, vol. 55, no. 8, pp. 3953–3964, Dec. 2023.

[59] P. Taylor, "The Growing Electoral Clout of Blacks Is Driven by Turnout, Not Demographics," Pew Research Center, Tech. Rep., 2013.

[60] US Census Bureau, *Current Population Survey Voting and Registration Supplements: 1980-2016*, 2017.

[61] B. Grofman, L. Handley, and R. G. Niemi, *Minority representation and the quest for voting equality*. Cambridge University Press, 1992.

[62] S. A. Banducci, T. Donovan, and J. A. Karp, "Minority Representation, Empowerment, and Participation," *The Journal of Politics*, vol. 66, no. 2, pp. 534–556, May 2004, Publisher: The University of Chicago Press.

[63] T. S. Arrington and G. L. Ingalls, “The limited vote alternative to affirmative districting,” *Political Geography*, vol. 17, no. 6, pp. 701–728, Aug. 1998.

[64] A. Lijphart and B. Grofman, Eds., *Choosing an electoral system: issues and alternatives* (American political parties and elections). Westport, Conn.: Praeger, 1984, ISBN: 978-0-03-069546-9 978-0-275-91216-1.

[65] S. Bowler, T. Donovan, and D. Brockington, *Electoral Reform and Minority Representation: Local Experiments with Alternative Elections*. Ohio State University Press, 2003, Google-Books-ID: 2LU520iehmgC, ISBN: 978-0-8142-0917-2.

[66] S. Bowler, T. Donovan, and D. M. Farrell, “Party Strategy and Voter Organization under Cumulative Voting in Victorian England,” *Political Studies*, vol. 47, no. 5, pp. 906–917, Dec. 1999, Publisher: SAGE Publications Ltd.

[67] R. L. Engstrom, “Cumulative and Limited Voting: Minority Electoral Opportunities and More,” . *Louis U. Pub. L. Rev.*, vol. 30, p. 97, 2010, Publisher: HeinOnline.

[68] R. H. Pildes and K. A. Donoghue, “Cumulative voting in the United States,” *U. Chi. Legal F.*, p. 241, 1995, Publisher: HeinOnline.

[69] L. Guinier, *Tyranny of the Majority: Funamental Fairness in Representative Democracy*. Free Press, 1994, Google-Books-ID: aRF7XdcCLq0C, ISBN: 978-0-02-913169-5.

[70] D. Brockington, T. Donovan, S. Bowler, and R. Brischetto, “Minority Representation under Cumulative and Limited Voting,” *The Journal of Politics*, vol. 60, no. 4, pp. 1108–1125, Nov. 1998, Publisher: The University of Chicago Press.

[71] S. Bowler, D. Brockington, and T. Donovan, “Election Systems and Voter Turnout: Experiments in the United States,” *Journal of Politics*, vol. 63, no. 3, pp. 902–915, 2001, https://onlinelibrary.wiley.com/doi/pdf/10.1111/0022-3816.00093.

[72] T. R. Palfrey and H. Rosenthal, “Voter Participation and Strategic Uncertainty,” *American Political Science Review*, vol. 79, no. 1, pp. 62–78, Mar. 1985.

[73] D. K. Levine and T. R. Palfrey, “The Paradox of Voter Participation? A Laboratory Study,” *American Political Science Review*, vol. 101, no. 1, pp. 143–158, Feb. 2007.

[74] L. Bursztyn, D. Cantoni, P. Funk, F. Schönenberger, and N. Yuchtman, *Identifying the Effect of Election Closeness on Voter Turnout: Evidence from Swiss Referenda*, Working Paper, Jun. 2017.

[75] E. R. Gerber, R. B. Morton, and T. A. Rietz, “Minority Representation in Multimember Districts,” *American Political Science Review*, vol. 92, no. 1, pp. 127–144, Mar. 1998.

[76] D. Cooper and A. Zillante, "A comparison of cumulative voting and generalized plurality voting," *Public Choice*, vol. 150, no. 1-2, pp. 363–383, Jan. 2012.

[77] P.P., "Report and Minutes of Evidence of the Select Committee on School Board Voting," Tech. Rep., 1884, p. 78.

[78] J. Sawyer and D. MacRae, "Game Theory and Cumulative Voting in Illinois: 1902–1954," *American Political Science Review*, vol. 56, no. 4, pp. 936–946, Dec. 1962.

[79] H. Herrera, M. Morelli, and T. Palfrey, "Turnout and power sharing," *The Economic Journal*, vol. 124, no. 574, F131–F162, 2014, Publisher: Oxford University Press Oxford, UK.

[80] M. Kartal, "Laboratory elections with endogenous turnout: Proportional representation versus majoritarian rule," *Experimental Economics*, vol. 18, no. 3, pp. 366–384, Sep. 2015.

[81] IGPA (Institute for Government and Public Affairs), "Illinois Assembly On Political Representation And Alternative Electoral Systems," University of Illinois, Tech. Rep., 2001.

[82] A. T. Cole, "Legal and Mathematical Aspects of Cumulative Voting," *South Carolina Law Review*, vol. 2, no. 3, p. 5, 1950.

[83] G. J. Glasser, "Game Theory and Cumulative Voting for Corporate Directors," *Management Science*, vol. 5, no. 2, pp. 151–156, Jan. 1959, Publisher: INFORMS.

[84] S. J. Brams, *Game Theory and Politics*. New York: Free Press, 1975, Google-Books-ID: 3cgWAAAAIAAJ, ISBN: 978-0-02-904550-3.

[85] A. Glazer, D. G. Glazer, and B. Grofman, "Cumulative Voting in Corporate Elections: Introducing Strategy into the Equation," *South Carolina Law Review*, vol. 35, no. 2, p. 3, 1984.

[86] M. Arzumanyan and M. K. Polborn, "Costly voting with multiple candidates under plurality rule," *Games and Economic Behavior*, vol. 106, pp. 38–50, Nov. 2017.

[87] D. K. Levine and A. Mattozzi, "Voter Turnout with Peer Punishment," *American Economic Review*, vol. 110, no. 10, pp. 3298–3314, Oct. 2020.

[88] C. B. Goldburg, "The Accuracy of Game Theory Predictions for Political Behavior: Cumulative Voting in Illinois Revisited," *The Journal of Politics*, vol. 56, no. 4, pp. 885–900, Nov. 1994, Publisher: The University of Chicago Press.

[89] A. Casella and A. Gelman, "A simple scheme to improve the efficiency of referenda," *Journal of Public Economics*, vol. 92, no. 10, pp. 2240–2261, Oct. 2008.

[90] U. Fischbacher, "Z-Tree: Zurich toolbox for ready-made economic experiments," *Experimental Economics*, vol. 10, no. 2, pp. 171–178, Jun. 2007.

[91] M. L. Duch, M. R. P. Grossmann, and T. Lauer, "*z-Tree unleashed*: A novel client-integrating architecture for conducting z-Tree experiments over the Internet," *Journal of Behavioral and Experimental Finance*, vol. 28, p. 100 400, Dec. 2020.

[92] A. Casella, A. Gelman, and T. R. Palfrey, "An experimental study of storable votes," *Games and Economic Behavior*, vol. 57, no. 1, pp. 123–154, Oct. 2006.

[93] J. Andreoni, Y.-K. Che, and J. Kim, "Asymmetric information about rivals' types in standard auctions: An experiment," *Games and Economic Behavior*, vol. 59, no. 2, pp. 240–259, May 2007.

[94] C. R. Taylor and H. Yildirim, "Public information and electoral bias," *Games and Economic Behavior*, vol. 68, no. 1, pp. 353–375, Jan. 2010.

[95] J. O. Ledyard, "The pure theory of large two-candidate elections," *Public Choice*, vol. 44, no. 1, pp. 7–41, Jan. 1984.

[96] C. R. Taylor and H. Yildirim, "A unified analysis of rational voting with private values and group-specific costs," *Games and Economic Behavior*, vol. 70, no. 2, pp. 457–471, Nov. 2010.

[97] V. Krishna and J. Morgan, "Majority Rule and Utilitarian Welfare," *American Economic Journal: Microeconomics*, vol. 7, no. 4, pp. 339–375, Nov. 2015.

Appendix A: Appendix for Chapter 1

A.1 Image Concerns

Recall that all agents have a utility function of the following form: $U_i(g_i, g_{-i}) = \pi_i(g_i, g_{-i}) + (1+b)W(g_i) + bpR(g_i)$. The following Proposition characterizes the Nash equilibrium contribution profile:

Proposition. The Nash equilibrium contribution profile of the one-shot game is given by $(g_1^*, ..., g_N^*)$, where for $i \in \{1, ..., N\}$:

$$g_i^* = \begin{cases} 0 & \text{if } 1 - \frac{m}{N} > \frac{\partial W}{\partial g_i} + b\frac{\partial W}{\partial g_i} + bp\frac{\partial R}{\partial g_i} \ \forall g_i \in [0, Y] \\ g_i & \text{such that } 1 - \frac{m}{N} = \frac{\partial W}{\partial g_i} + b\frac{\partial W}{\partial g_i} + bp\frac{\partial R}{\partial g_i} \\ Y & \text{if } 1 - \frac{m}{N} < \frac{\partial W}{\partial g_i} + b\frac{\partial W}{\partial g_i} + bp\frac{\partial R}{\partial g_i} \ \forall g_i \in [0, Y] \end{cases} \tag{A.1}$$

Additionally, if all agents are symmetric, then $g_1^* = ... = g_N^*$.

Proof. The condition in the expression for g_i^* comes from the first-order condition of the agent's utility maximization problem. On the left-hand side is the marginal payoff from not contributing a marginal unit to the public good: $1 - \frac{m}{N}$. On the right-hand side is the image utility gain from contributing a marginal unit to the public good: $\frac{\partial W}{\partial g_i} + b\frac{\partial W}{\partial g_i} + bp\frac{\partial R}{\partial g_i}$. If interior, g_i^* equates the two sides. Edge cases occur if one side is always larger than the other: if the marginal payoff from not contributing always exceeds the image gain from contributing, then $g_i^* = 0$ is the utility-maximizing contribution. A similar argument justifies when $g_i^* = Y$.

A.2 Punishment outside the Game

The punishment game is played with N players indexed by $i \in \{1, ..., N\}$. The action space is given by $A = \{P, X\}^N$. Each player's action profile is a N-dimensional vector $a_i = (a_{i1}, a_{i2}, ..., a_{iN})$, where $a_{ij} \in \{P, X\}$ denotes whether player i punishes (P) or does not punish (X) player j. There is a cost $c > 0$ to punish another player, as well as a cost $C >> c$ if punished by another player. The payoff in the stage game is given by $\pi_i = -c \sum_{j=1}^{N} \mathbb{I}(a_{ij} = P) - C \sum_{j=1}^{N} \mathbb{I}(a_{ji} = P)$.[1]

The punishment game is repeated infinitely in discrete time, with discount factor $\delta \in (0, 1)$. I assume there is perfect monitoring in the punishment game: at the end of each period, all players observe each other's action profiles. In other words, all players know who punishes who in each period. The following Proposition characterizes the relationship between the level of anonymity in a public goods game and the expected punishment from free riding in the subsequent punishment game. The intuition is straightforward: when a deviator's identity can be known, the other group members are able to coordinate punishment on the deviator. When the deviator's identity cannot be known, the deviator escapes punishment with strictly positive probability. Hence, the expected punishment from deviating is strictly lower.

Proposition. Let a finitely repeated public goods game be followed by an infinitely repeated punishment game. There exists some discount factor $\delta \in (0, 1)$ that can sustain positive contributions in the public goods game (1) a deviation can be attributed with certainty to an identifier that (2) allows the deviator to be identified outside the game, but cannot sustain positive contributions otherwise. The reverse does not hold.

Proof. I begin by restricting focus to the scenario when it is common knowledge that there exists a "target" for punishment.[2] The proof then proceeds in two parts. First, I consider the case when the target's identity is also commonly known. For this case, I construct a subgame perfect

[1] The unique stage game Nash equilibrium is given by $a_{ij} = X$ for all i, j. Nobody punishes, and all players earn payoff $\pi = 0$. This is because punishment is costly to administer, and to incur.

[2] Given some strategy in the public goods game, players can deduce just from the aggregate contribution to the public good, whether a group member deviated. In all experimental conditions, the aggregate contribution is made known after every round.

Nash equilibrium in which the target is punished indefinitely by all other group members. Second, I consider the case where the target's identity cannot be commonly known. I show that in any equilibrium, the target escapes punishment with strictly positive probability.

I assume in the proof, without loss of generality, that the target is Player 1.

Part 1. Suppose that it is common knowledge among all players that Player 1 is the target: Player 1's actions were attributed with certainty to an identifier during the public goods game that allowed their group members to identify them during the punishment game. Consider the following set of strategies in which punishment is coordinated on the target:

$$\begin{cases} a_1 & = (X, X, ..., X) \\ a_i & = (P, X, ..., X), \ \forall i \in \{2, ..., N\} \end{cases}$$

This set of strategies can be sustained as a subgame perfect Nash equilibrium in the infinitely repeated game, so long as there is no profitable one-shot deviation.

It is straightforward to see that Player 1 (the target) has no profitable deviation from their strategy.

Without loss of generality, consider whether Player 2 has a profitable deviation. Suppose that Player 2 deviates from their strategy, and does not punish the target in a given period. In that period, Player 2 has a payoff of $0 > -c$. But starting in the following period, all other players $i \in \{1, 3, ..., N\}$ can punish Player 2 forever for deviating; in effect, Player 2 would become the new target. Therefore, Player 2 does not have a profitable deviation if: $-\frac{c}{1-\delta} \geq 0 - \frac{\delta}{1-\delta}(N-1)C$. Rearranging terms yields: $\delta \geq \frac{c}{(N-1)C}$. Because $\frac{c}{(N-1)C} < 1$, this set of strategies that coordinates punishment on the target can be sustained as a subgame perfect Nash equilibrium in the infinitely repeated game.

Note that when the target's identity is common knowledge, the target receives payoff $\pi_1 = -(N-1)C$ each period.

Part 2. Suppose the target's identity is not common knowledge. The only way in which the players can guarantee that the target receives payoff $-(N-1)C$ per period is if everyone punishes everyone else: $a_{ij} = P$, $\forall i \neq j$. However, this cannot be sustained in equilibrium: a player may profitably deviate by not punishing, and they cannot be punished for the deviation since they are already being punished by everyone else. Therefore, in any equilibrium of the game when the target's identity is not common knowledge, the target's expected payoff is strictly greater (*i.e.*, less negative) than $-(N-1)C$.

Denote the expected punishment from free riding as $-P^{K}_{FR} = -(N-1)C$ when the deviator's identity is commonly known and as $-P^{UK}_{FR}$ when the deviator's identity cannot be commonly known. The expected punishment from free riding in the public goods game is strictly smaller when there is sufficient anonymity to mask the deviator's identity: $-P^{K}_{FR} < -P^{UK}_{FR}$. Hence, the minimum discount rate δ that can support positive contributions satisfies: $\delta^{K} < \delta^{UK}$. Therefore, there exists a δ that supports positive contributions when the deviator's identity can be commonly known that cannot support positive contributions when the deviator's identity cannot be commonly known.

Appendix B: Appendix for Chapter 3

B.1 CV when voting is costless

Proposition 1. *In the absence of voting costs, in all party-optimal equilibria of the CV voting game: (i) for all $m < M/K$, the minority never wins any seat; (ii) for all $m \geq M/K$:*

$$z = \left\lfloor \frac{Km+m}{M+m} \right\rfloor \qquad \textit{if } \frac{Km+m}{M+m} \notin \mathbb{Z}$$

$$z = \begin{cases} \frac{Km+m}{M+m} - 1 & \text{with prob } m/(m+M) \\ \frac{Km+m}{M+m} & \text{with prob } M/(m+M) \end{cases} \qquad \textit{if } \frac{Km+m}{M+m} \in \mathbb{Z}$$

Proof. We establish the proposition by proceeding in three steps.

1. First, we note that the identity of purpose between party leaders and voters implies that in all party-optimal equilibria we can think of the party leaders as controlling not only the number of party candidates but also the distribution of votes cast by their party voters.

2. Second, we show that the proposition identifies the number of seats won by the minority when both parties follow maximin strategies.

(i) Suppose first that $m < M/K$. Then the M party can guarantee itself all K seats by dividing its votes equally over K candidates, and the m party cannot win any seat.

(ii) Suppose then $m > M/K$. For any x_{M}, party m maximizes the probability of winning z seats by dividing its votes equally over z candidates, and guarantees itself z seats if $mK/z > MK/(K-z+1)$, or $z < (Km+m)/(M+m)$. At the same time, party M maximizes the probability of winning $(K-z)$ seats by dividing its votes equally over $K-z$ candidates, and guarantees itself $K-z$ seats if $MK/(K-z) > mK/(z+1)$, or $z > (Km-M)/(M+m)$. We require z to be an integer.

Note that $(Km - M)/(M + m) = (Km + m)/(M + m) - 1$. Hence either both $(Km + m)/(M + m)$ and $(Km - M)/(M + m)$ are integers, or neither one is an integer.

(ii.a) Suppose first that $(Km + m)/(M + m)$ is not an integer. Party m guarantees itself $\lfloor \frac{Km+m}{M+m} \rfloor$seats, and party M guarantees itself $K - \lfloor \frac{Km+m}{M+m} \rfloor$seats.

(ii.b and iii) Finally, suppose that either $m > M/K$ and $(Km + m)/(M + m)$ is an integer, or $m = M/K$ (and thus $(Km + m)/(M + m) = 1$). Then the m party can guarantee itself $(Km + m)/(M+m)-1 \equiv \underline{z}$ seats, but can do better by spreading votes equally over $(Km+m)/(M+m) \equiv \overline{z}$ candidates. Similarly, the M party can guarantee itself $K - \overline{z}$ seats, but can do better by spreading votes equally over $K - \overline{z} + 1 = K - \underline{z}$ candidates. In equilibrium then, party m (M) spreads its votes equally over $\overline{z}$ ($K - \underline{z}$) candidates; a total of $K + 1$ candidates receive votes, and all are tied with $[K/(K + 1)](M + m)$ votes each. The tie-break rule selects K winners randomly from the $K + 1$ candidates. It then follows that:

$$prob(z = \overline{z}) = \frac{\binom{K - \underline{z}}{K - \overline{z}}}{\binom{K+1}{K}} = 1 - \frac{\overline{z}}{K + 1} = \frac{M}{m + M}$$

$$prob(z = \underline{z}) = 1 - prob(z = \overline{z}) = \frac{m}{m + M}$$

3. Because $u(k)$ is linear in k, the game is constant-sum. It follows that party-optimal equilibria are equilibria of a constant sum, two-player game. Hence, if equilibria exist, they must all yield maximin payoffs. It is not difficult to verify that the strategies described above are equilibria: neither party has a profitable deviation. Thus equilibria exist and all yield z minority victories. □

We complement the proof with two observations. First, as noted in the text, for given K, m, and M, in general multiple party-optimal equilibria exist, with different numbers of candidates and/or distributions of votes. For example, suppose $K = 4$, $m = 3$, $M = 6$, and describe an equilibrium by a vector $\{g, G, \{x^k_{\mathsf{m}}\}, \{x^k_{\mathsf{M}}\}\}$. Then $\{1, 3, \{12\}, \{8, 8, 8\}\}$ is an equilibrium. But so are $\{1, 4, \{12\}, \{6, 6, 6, 6\}\}$; $\{1, 3, \{12\}, \{10, 7, 7\}\}$; $\{1, 3, \{12\}, \{9, 8, 7\}\}$; $\{2, 3, \{8, 4\}, \{8, 8, 8\}\}$; $\{2, 3, \{8, 4\}, \{10, 7, 7\}\}$; $\{1, 3, \{8, 4\}, \{9, 8, 7\}\}$, and there are many others. In all party-optimal equilibria, however, as the proposition states, $z = 1$ in this example.[1]

Second, to clarify the logic of the CV game, it is useful to differentiate it from a Colonel Blotto game, adapted to the parameters used here. In the Blotto game, two players, with Km and KM tokens respectively, simultaneously distribute them over K boxes; each player earns one point for each box in which the player's tokens are more numerous than the opponent's. In the CV game, each of the two players, again endowed with Km and KM tokens respectively, has a separate set of K boxes over which to distribute the tokens; the K boxes with most tokens are chosen, out of the total $2K$ boxes, and each player earns 1 point for each box chosen out of the player's own set of K. The two games are different. For example, in the Blotto game, the equilibrium typically requires mixed strategies, and the player with fewer tokens cannot be guaranteed any points; neither statement applies to the CV game.

B.2 Costly voting

B.2.1 Multi-winner plurality (MP). Pivot probabilities and probabilities of winning seats

We report here the binomial formulas for the pivot probabilities. Under MP, such formulas are well-known (see for example [73]).

$$\pi_{\mathsf{m}}^{T-1} = \sum_{x=0}^{m-1} \binom{m-1}{x} \binom{M}{x+1} F(c_{\mathsf{m}})^x [1 - F(c_{\mathsf{m}})]^{m-1-x} F(c_{\mathsf{M}})^{x+1} [1 - F(c_{\mathsf{M}})]^{M-(x+1)}$$

$$\pi_{\mathsf{m}}^{T} = \sum_{x=0}^{m-1} \binom{m-1}{x} \binom{M}{x} F(c_{\mathsf{m}})^x [1 - F(c_{\mathsf{m}})]^{m-1-x} F(c_{\mathsf{M}})^{x} [1 - F(c_{\mathsf{M}})]^{M-x}$$

[1] Other equilibria exist that are not party-optimal, where the lack of coordination by the voters of one of the parties prevents it from winning all the seats it could win. In the example above, $\{2, 3, \{6, 6\}, \{12, 12, 0\}\}$ is an equilibrium: majority voters fail to coordinate and because each only holds 4 votes, no profitable individual deviation exists. Each party wins two seats.

and:

$$\pi_{\mathsf{M}}^{T-1} = \sum_{x=1}^{m} \binom{m}{x}\binom{M-1}{x-1} F(c_{\mathsf{m}})^x [1 - F(c_{\mathsf{m}})]^{m-x} F(c_{\mathsf{M}})^{x-1} [1 - F(c_{\mathsf{M}})]^{M-1-(x-1)}$$
$$\pi_{\mathsf{M}}^{T} = \sum_{x=0}^{m} \binom{m}{x}\binom{M-1}{x} F(c_{\mathsf{m}})^x [1 - F(c_{\mathsf{m}})]^{m-x} F(c_{\mathsf{M}})^{x} [1 - F(c_{\mathsf{M}})]^{M-1-x}$$

The frequency of minority victories is sensitive to the relative turnout rates of the two parties, captured by the two thresholds c_{m} and c_{M}. Although the study of costly voting models has identified an "underdog effect"—the tendency for the minority's turnout rate to be higher than the majority's, or $c_{\mathsf{m}} > c_{\mathsf{M}}$—the existence of such an effect is sensitive to the exact specification of the model. It has been proven in a number of scenarios: when the voting cost is fixed and equal for all ([94]); when voters' direction of preferences is randomly drawn ([95]; [96]); when the size of the electorate is uncertain ([79]; [97]). The specification used here differs from these models, and relative turnout under MP depends on V, the value of winning all seats. Because the model is widely used but this observation is missing from the literature, we make it explicit in the following remark.

Remark. *For any finite $M \geq m$, and F continuous and atomless over support $[\underline{c}, \overline{c}]$,with $\underline{c} \geq 0$, there exists a finite $\widehat{V}(M, m)$ such that if $V = \widehat{V}$, then there exists an equilibrium with $c_{\mathsf{m}} = c_{\mathsf{M}}$.*

Proof. Call $\widehat{c}$ the median of $F(c)$. Straightforward manipulations of the pivot probabilities show that if $c_{\mathsf{m}} = c_{\mathsf{M}} = \widehat{c}$, and thus $F(c_{\mathsf{m}}) = 1 - F(c_{\mathsf{M}}) = 1/2$, then $(\pi_{\mathsf{m}}^T + \pi_{\mathsf{m}}^{T-1}) = (\pi_{\mathsf{M}}^T + \pi_{\mathsf{M}}^{T-1}) = (1/2)^{M+m-1}\binom{M+m}{m}$. Hence for any M and m, $c_{\mathsf{m}} = c_{\mathsf{M}} = \widehat{c}$ is an equilibrium as long as $\widehat{c} = (V/2)(1/2)^{M+m-1}\binom{M+m}{m}$, or $V = \widehat{c}\left(2^{(M+m)}/\binom{M+m}{m}\right) = \widehat{V}$. $\square$

The derivation of the probabilities of winning different numbers of seats is straightforward. Consider the problem from the perspective of a minority voter. Begin with the probability of losing all positions, $\Pr(W_{\mathsf{m}} = 0)$. Such probability equals the probability that either all minority candidates receive strictly fewer votes than the majority candidates, or that all candidates are tied but minority candidates lose all tie-breaks. Or, $\Pr(W_{\mathsf{m}} = 0) = \Pr(S_{\mathsf{m}} < S_{\mathsf{M}}) + \Pr[(S_{\mathsf{m}} = S_{\mathsf{M}}) \cap$

(m loses all tie-breaks)]. That is:

$$\Pr(W_{\mathsf{m}} = 0) =$$
$$= \sum_{S_{\mathsf{M}}=1}^{M} \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \sum_{S_{\mathsf{m}}=0}^{S_{\mathsf{M}}-1} \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}}[1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{m}}} +$$
$$+ \sum_{S_{\mathsf{M}}=0}^{M} \binom{M}{S_{\mathsf{M}}}\binom{m}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \times$$
$$\times F(c_{\mathsf{m}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{M}}} \left(1/\binom{2K}{K}\right)$$

Similarly, the probability that m wins all positions, $\Pr(W_{\mathsf{m}} = K)$ equals the probability that either all minority candidates receive strictly more votes than the majority candidates, or that all candidates are tied but minority candidates win all tie-breaks. Or, $\Pr(W_{\mathsf{m}} = K) = \Pr(S_{\mathsf{m}} > S_{\mathsf{M}}) + \Pr[(S_{\mathsf{m}} = S_{\mathsf{M}}) \cap$ (m wins all tie-breaks)]. That is:

$$\Pr(W_{\mathsf{m}} = K) =$$
$$= \sum_{S_{\mathsf{M}}=0}^{m-1} \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \sum_{S_{\mathsf{m}}=S_{\mathsf{M}}+1}^{m} \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}}[1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{m}}} +$$
$$+ \sum_{S_{\mathsf{M}}=0}^{M} \binom{M}{S_{\mathsf{M}}}\binom{m}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \times$$
$$\times F(c_{\mathsf{m}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{M}}} \left(1/\binom{2K}{K}\right)$$

The probabilities of other numbers of minority victories can be derived in the same fashion. The probability of electing w minority candidates, with $w \in (0, K)$ equals the probability that all candidates are tied and m wins w tie-breaks. Thus:

$$\begin{aligned}
&\Pr(W_{\mathsf{m}} = w) = \\
&\quad = \sum_{S_{\mathsf{M}}=0}^{M} \binom{m}{S_{\mathsf{M}}} F(c_{\mathsf{m}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{m}})]^{m - S_{\mathsf{M}}} \times \\
&\quad \times \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{M}})]^{M - S_{\mathsf{M}}} \binom{K}{w} \binom{K}{K - w} / \binom{2K}{K}
\end{aligned}$$

For given M, m, K, $F(c)$, and $\{u(k)\}$, for $k \in \{0, 1, .., K\}$, the equilibrium yields expected turnout rates for voters of the two parties, the probabilities of winning $0, 1, .., K$ positions for each party, and ex ante expected utility for an M and an m voter.[2]

B.2.2 Cumulative Voting (CV). Pivot probabilities and probabilities of winning seats

Consider first the perspective of a majority voter. The pivot probabilities correspond to the probabilities of the three events described in the text: breaking a tie (if $(K/G)S_{\mathsf{M-i}} = (K/g)S_{\mathsf{m}}$), making a tie (if $(K/G)(S_{\mathsf{M-i}} + 1) = (K/g)S_{\mathsf{m}}$), or moving the outcome from a loss to a win on all contested positions (if $S_{\mathsf{M-i}} \in (S_{\mathsf{m}}(G/g) - 1, S_{\mathsf{m}}(G/g))$). Note that since $S_{\mathsf{M}-i}$ and S_{m} are non-negative integers, the first event is only possible if either G/g is an integer, or $S_{\mathsf{M}-i} = S_{\mathsf{m}} = 0$; the second event is only possible if G/g is an integer, and the third event is only possible if G/g is not an integer.

The equations corresponding to the pivot probabilities are logically straightforward:

$$\begin{aligned}
\widetilde{\pi}_{\mathsf{M}}^{T} = I_Q[(G/g)S_{\mathsf{m}}] \sum_{S_{\mathsf{m}}=0}^{m} \Bigg\{ & \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}} [1 - F(c_{\mathsf{m}})]^{m - S_{\mathsf{m}}} \\
& \binom{M - 1}{(G/g)S_{\mathsf{m}}} F(c_{\mathsf{M}})^{(G/g)S_{\mathsf{m}}} [1 - F(c_{\mathsf{M}})]^{M - 1 - (G/g)S_{\mathsf{m}}} \Bigg\}
\end{aligned}$$

[2] Note in particular that if $u(K) - Eu_T^{MP} = Eu_T^{MP} - u(0)$, or $Eu_T^{MP} = [u(K) - u(0)]/2$, the equilibrium thresholds $\{c_{\mathsf{m}}, c_{\mathsf{M}}\}$ are identical to the thresholds that solve the corresponding costly voting problem with a single winner.

$$\widetilde{\pi}_{\mathsf{M}}^{T-1} = I_Q[(G/g)S_{\mathsf{m}}] \sum_{S_{\mathsf{m}}=1}^{m} \left\{ \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}} [1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{m}}} \right.$$
$$\left. \binom{M-1}{(G/g)S_{\mathsf{m}}-1} F(c_{\mathsf{M}})^{(G/g)S_{\mathsf{m}}-1} [1 - F(c_{\mathsf{M}})]^{M-1-[(G/g)S_{\mathsf{m}}-1]} \right\}$$

and

$$\widetilde{\pi}_{\mathsf{M}}^{W} = \left(1 - I_Q[(G/g)S_{\mathsf{m}}]\right) \sum_{S_{\mathsf{m}}=0}^{m} \left\{ \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}} [1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{m}}} \right.$$
$$\left. \binom{M-1}{\lfloor (G/g)S_{\mathsf{m}} \rfloor} F(c_{\mathsf{M}})^{\lfloor (G/g)S_{\mathsf{m}} \rfloor} [1 - F(c_{\mathsf{M}})]^{M-1-\lfloor (G/g)S_{\mathsf{m}} \rfloor} \right\}$$

where $I_Q[(G/g)S_{\mathsf{m}}] = 1$ if $(G/g)S_{\mathsf{m}}$ is an integer, and 0 otherwise, and $\lfloor x \rfloor$ is the floor function, denoting the greatest integer smaller or equal to x.[3]

The problem is analogous for a minority voter. The relevant equations are:

$$\widetilde{\pi}_{\mathsf{m}}^{T} = I_Q[(g/G)S_{\mathsf{M}}] \left\{ \sum_{S_{\mathsf{M}}=0}^{M} \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \right.$$
$$\left. \binom{m-1}{(g/G)S_{\mathsf{M}}} F(c_{\mathsf{m}})^{(g/G)S_{\mathsf{M}}} [1 - F(c_{\mathsf{m}})]^{m-1-(g/G)S_{\mathsf{M}}} \right\}$$

$$\widetilde{\pi}_{\mathsf{m}}^{T-1} = I_Q[(g/G)S_{\mathsf{M}}] \sum_{S_{\mathsf{M}}=1}^{M} \left\{ \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \right.$$
$$\left. \binom{m-1}{(g/G)S_{\mathsf{M}}-1} F(c_{\mathsf{m}})^{(g/G)S_{\mathsf{M}}-1} [1 - F(c_{\mathsf{m}})]^{m-1-[(g/G)S_{\mathsf{M}}-1]} \right\}$$

$$\widetilde{\pi}_{\mathsf{m}}^{W} = \left(1 - I_Q[(g/G)S_{\mathsf{M}}]\right) \sum_{S_{\mathsf{M}}=0}^{M} \left\{ \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \right.$$
$$\left. \binom{m-1}{\lfloor (g/G)S_{\mathsf{M}} \rfloor} F(c_{\mathsf{m}})^{\lfloor (g/G)S_{\mathsf{M}} \rfloor} [1 - F(c_{\mathsf{m}})]^{m-1-\lfloor (g/G)S_{\mathsf{M}} \rfloor} \right\}$$

The probabilities of the minority winning different numbers of positions can be derived as

[3] Recall that $\binom{r}{y} = 0$ if $y > r$.

under MP, but taking into account that the number of candidates, in each party, now may differ from the number of seats. The probability of the minority losing all seats must be 0 if $G < K$; if instead $G \geq K$, then as before it equals the probability that either all minority candidates receive strictly lower votes than the majority candidates, or that all candidates are tied but minority candidates lose all tie-breaks. That is:

$$\begin{aligned}
&\Pr(W_{\mathsf{m}} = 0 | G \geq K) = \\
&= \sum_{S_{\mathsf{M}}=1}^{M} \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{M}})]^{M - S_{\mathsf{M}}} \sum_{S_{\mathsf{m}}=0}^{X(S_{\mathsf{M}})} \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}} [1 - F(c_{\mathsf{m}})]^{m - S_{\mathsf{m}}} + \\
&+ \sum_{S_{\mathsf{M}}=0}^{M} \binom{M}{S_{\mathsf{M}}} \binom{m}{(g/G) S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}} [1 - F(c_{\mathsf{M}})]^{M - S_{\mathsf{M}}} \times \\
&\times F(c_{\mathsf{m}})^{(g/G) S_{\mathsf{M}}} [1 - F(c_{\mathsf{m}})]^{m - (g/G) S_{\mathsf{M}}} I_Q[(g/G) S_{\mathsf{M}}] \binom{G}{K} / \binom{G+g}{K}
\end{aligned}$$

where:

$$X(S_{\mathsf{M}}) = \begin{cases} (g/G) S_{\mathsf{M}} - 1 & \text{if } (g/G) S_{\mathsf{M}} \text{ is an integer} \\ \lfloor (g/G) S_{\mathsf{M}} \rfloor & \text{otherwise} \end{cases}$$

The probability of electing some but not all minority candidates can be derived analogously. For any $w \in (0, g)$, the probability of electing w minority candidates is 0 if $K - G > w$; it equals the probability that all candidates are tied and m wins w tie-breaks if $K - G < w$, and equals the probability either that all are tied and m loses all tie-breaks or that all m candidates receive fewer votes if $K - G = w$. Thus:

$$
\begin{aligned}
&\Pr(W_{\mathsf{m}} = w | K - G \leq w) = \\
&\quad = \sum_{S_{\mathsf{M}}=0}^{M} \binom{m}{(g/G)S_{\mathsf{M}}} F(c_{\mathsf{m}})^{(g/G)S_{\mathsf{M}}}[1 - F(c_{\mathsf{m}})]^{m-(g/G)S_{\mathsf{M}}} \times \\
&\quad \times \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} I_Q[(g/G)S_{\mathsf{M}}] \binom{g}{w}\binom{G}{K-w} / \binom{G+g}{K} + \\
&\quad + I_{K-G=w} \sum_{S_{\mathsf{M}}=1}^{M} \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \times \\
&\quad \left(\sum_{S_{\mathsf{m}}=0}^{X(S_{\mathsf{M}})} \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}}[1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{m}}} \right)
\end{aligned}
$$

where $I_Q[(g/G)S_{\mathsf{M}}]$ and $X(S_{\mathsf{M}})$ are defined as above, and $I_{K-G=w}$ is an indicator function taking value 1 if $K - G = w$ and 0 otherwise.

Finally, the probability of electing exactly g minority candidates equals 1 if $K - G \geq g$; it equals the probability that either all minority candidates receive more votes or that all candidates are tied and the g minority candidates win all tie-breaks. That is:

$$
\begin{aligned}
&\Pr(W_{\mathsf{m}} = g | K - G < g) = \\
&\quad = \sum_{S_{\mathsf{M}}=0}^{M} \binom{m}{(g/G)S_{\mathsf{M}}} F(c_{\mathsf{m}})^{(g/G)S_{\mathsf{M}}}[1 - F(c_{\mathsf{m}})]^{m-(g/G)S_{\mathsf{M}}} \times \\
&\quad \times \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} I_Q[(g/G)S_{\mathsf{M}}] \binom{G}{K-g} / \binom{G+g}{K} + \\
&\quad \sum_{S_{\mathsf{m}}=1}^{m} \binom{m}{S_{\mathsf{m}}} F(c_{\mathsf{m}})^{S_{\mathsf{m}}}[1 - F(c_{\mathsf{m}})]^{m-S_{\mathsf{m}}} \times \\
&\quad \left(\sum_{S_{\mathsf{M}}=0}^{Y(S_{\mathsf{m}})} \binom{M}{S_{\mathsf{M}}} F(c_{\mathsf{M}})^{S_{\mathsf{M}}}[1 - F(c_{\mathsf{M}})]^{M-S_{\mathsf{M}}} \right)
\end{aligned}
$$

where:

$$
Y(S_{\mathsf{m}}) = \begin{cases} (G/g)S_{\mathsf{m}} - 1 & \text{if } (G/g)S_{\mathsf{m}} \text{ is an integer} \\ \lfloor (G/g)S_{\mathsf{m}} \rfloor & \text{otherwise} \end{cases}
$$

B.3 Additional Experimental Results

B.3.1 Individual turnout decisions

We tested the robustness of our results by estimating the regression in Table 3.4 while varying the exact specification or procedure. We report the results below.

Between-subjects design: first treatments only

Although we ran the experiment with a within-subjects design—i.e., with each subject playing multiple treatments—varying the order of treatments across sessions allows us to check the robustness of the results to a between-subjects design—i.e., if each subject only played one treatment. We can study the data from the first treatment played in each session. Table B.1 reproduces the regression in Table 3.4 in the text, restricting attention to such data. Some estimates lose precision, but the results are fully consistent with those obtained with the full set of data. In particular, the parameter of the interaction term remains positive in all treatments and statistically significant if $K = 2$.

Probit model

The turnout regression reported in Table B.1 in the text estimates a linear probability model. Alternatively, we can estimate a probit model, and the comparison of the results constitutes a second robustness check. As Table B.2 shows, here too the parameter of the interaction term remains positive in all treatments and statistically significant if $K = 2$. [4]

Separating the effect of treatment order on CV

The regression reported in Table B.3 adds as an explanatory variable the interaction term CV×CVfirst, allowing the effect of treatment order to affect CV and MP differently. As the table shows, we find no effect.

[4] The numbers reported in the table are raw coefficients from the probit regression.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
Minority	-0.437***	-0.072	-0.132*	-0.119
	(0.074)	(0.099)	(0.071)	(0.088)
CV	-0.062	0.077	-0.153**	-0.141
	(0.087)	(0.083)	(0.073)	(0.091)
CV × Minority	0.347**	0.048	0.235**	0.158
	(0.141)	(0.131)	(0.108)	(0.125)
Voting Cost	-0.007***	-0.009***	-0.009***	-0.007***
	(0.001)	(0.001)	(0.001)	(0.001)
Round	-0.006	-0.004	-0.004	-0.001
	(0.004)	(0.004)	(0.003)	(0.004)
Constant	1.116***	1.086***	1.204***	1.149***
	(0.055)	(0.073)	(0.052)	(0.060)
Observations	720	720	840	840
R^2	0.304	0.289	0.312	0.192

Standard errors in parentheses. 48 clusters in columns 1,2; 56 in columns 3,4.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.1: *Individual turnout decisions*. Data restricted to treatments run first in each session. Standard errors are clustered at the individual level.

Clustering

Our data consist of multiple observations for each subject. To allow for possible correlation among the subject's decisions, the standard errors reported in Table 3.4 in the text are clustered at the individual level. As mentioned in footnote 29, different levels of clustering can be defended. Because turnout is a strategic decision that depends on the other voters' expected turnout, standard errors could be clustered at the group level (i.e., at the level of each group of size $N = m + M$ disputing the same K seats). Results are reported in Table B.4.

Clustering at the group level strengthens our results by reducing standard errors, relative to Table 3.4 in the text, thus suggesting higher precision in the estimates of the parameters. This applies particularly to the parameters of the interaction term, capturing the difference-in-difference effect of CV on relative turnout. These results however should be evaluated with some caution. In our experimental design groups are rematched randomly after every round. The implication is that, from a substantive point of view, it is de facto impossible to form precise beliefs about other

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
Turnout				
Minority	-0.724***	-0.146	-0.538***	-0.722***
	(0.172)	(0.168)	(0.160)	(0.183)
CV	0.027	0.099	-0.466***	-0.089
	(0.117)	(0.118)	(0.137)	(0.117)
CV × Minority	0.707***	0.401	0.799***	0.316
	(0.255)	(0.246)	(0.243)	(0.216)
Voting Cost	-0.023***	-0.024***	-0.029***	-0.027***
	(0.002)	(0.002)	(0.002)	(0.002)
CVfirst	-0.182	-0.088	0.095	-0.022
	(0.153)	(0.156)	(0.140)	(0.147)
Round	-0.026***	-0.009	-0.028***	-0.020***
	(0.007)	(0.006)	(0.006)	(0.007)
Constant	1.857***	1.741***	2.412***	2.424***
	(0.158)	(0.193)	(0.151)	(0.176)
Observations	2880	2880	3360	3360
$pseudo-R^2$	0.196	0.193	0.256	0.242

Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, *** $p<0.01$

Table B.2: *Individual turnout decisions*. Estimated probit model.

group members' choices, while, from a statistical point of view, the same group rarely repeats. When clustering at the group level, the number of clusters approaches 400 in all parametrizations. Clustering at the individual level, as in Table 3.4, is more conservative.

At the other extreme, it is possible to argue for clustering at the session level, allowing for correlation among the choices of all subjects sharing the same session. Table B.5 reports the results.

The main effect is on the standard error of the parameter of the interaction term, increasing it for $m = 2$ treatments, and decreasing it if $m = 3$, thus suggesting positive correlation among errors in $m = 2$ sessions, but negative correlation in $m = 3$ sessions. Although possible, the divergence is not easy to rationalize. In committee voting experiments, the choice of a single voter can strongly affect the result, contrary, for example, to market experiments. As a result, individual experiences can differ strongly among participants to the same experimental session.

	(1)	(2)	(3)	(4)
	m2K2	m2K4	m3K2	m3K4
Minority	-0.229***	-0.042	-0.144***	-0.202***
	(0.052)	(0.052)	(0.045)	(0.051)
CV	0.042	0.046	-0.096**	0.009
	(0.049)	(0.043)	(0.046)	(0.039)
CV × Minority	0.227***	0.118	0.218***	0.091
	(0.078)	(0.072)	(0.069)	(0.061)
Voting Cost	-0.007***	-0.008***	-0.009***	-0.008***
	(0.001)	(0.001)	(0.000)	(0.001)
CVfirst	-0.023	-0.012	0.063	0.029
	(0.051)	(0.053)	(0.045)	(0.050)
CV × CVfirst	-0.067	-0.034	-0.069	-0.071
	(0.053)	(0.049)	(0.049)	(0.045)
Round	-0.007***	-0.003	-0.007***	-0.005**
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	1.075***	1.038***	1.180***	1.167***
	(0.041)	(0.051)	(0.035)	(0.042)
Observations	2880	2880	3360	3360
R^2	0.240	0.228	0.297	0.270

Standard errors in parentheses. 96 clusters in columns 1 and 2, 112 in columns 3 and 4.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.3: *Individual turnout decisions*. The CV×CVfirst term is added to the regression specification. Standard errors are clustered at the individual level.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
Minority	-0.229***	-0.042*	-0.144***	-0.202***
	(0.026)	(0.025)	(0.021)	(0.020)
CV	0.008	0.029	-0.130***	-0.027
	(0.021)	(0.018)	(0.019)	(0.018)
CV × Minority	0.227***	0.118***	0.218***	0.091***
	(0.040)	(0.035)	(0.029)	(0.028)
Voting Cost	-0.007***	-0.008***	-0.009***	-0.008***
	(0.000)	(0.000)	(0.000)	(0.000)
CVfirst	-0.057***	-0.029*	0.029*	-0.006
	(0.017)	(0.015)	(0.015)	(0.015)
Round	-0.008***	-0.003*	-0.008***	-0.005***
	(0.002)	(0.002)	(0.002)	(0.002)
Constant	1.095***	1.048***	1.201***	1.189***
	(0.023)	(0.023)	(0.020)	(0.019)
Observations	2880	2880	3360	3360
R^2	0.239	0.228	0.296	0.268

Standard errors in parentheses. 387 clusters in column 1, 414 in column 2, 448 in columns 3 and 4.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.4: *Individual turnout decisions*. Standard errors are clustered at the group level.

For this reason, and because results across different estimation methods are more consistent under individual level clustering than under session-level clustering, the turnout regression reported in the text clusters standard errors at the individual level.

Random effects

The estimation of individual cost cutpoints in Section 3.5.3 shows large dispersion, suggesting that controlling for individual fixed effects could be desirable. However, in our experimental design, individuals' assignment to either the minority or the majority party is random and changes across treatments.[5] Thus not all subjects experience both roles, and controlling for individual fixed effects would effectively drop from the estimation of the parameters of interest those subjects who happen to be assigned to the same party under both MP and CV. A better procedure then is to exploit the panel nature of the data set by using a random effects model, where unobserved het-

[5]Individuals maintain their identity as either minority or majority member throughout a treatment.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
Minority	-0.229**	-0.042	-0.144**	-0.202**
	(0.066)	(0.044)	(0.060)	(0.059)
CV	0.008	0.029	-0.130***	-0.027
	(0.037)	(0.030)	(0.031)	(0.028)
CV × Minority	0.227	0.118	0.218**	0.091*
	(0.151)	(0.098)	(0.065)	(0.041)
Voting Cost	-0.007***	-0.008***	-0.009***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.001)
CVfirst	-0.057	-0.029	0.029	-0.006
	(0.061)	(0.028)	(0.051)	(0.061)
Round	-0.008**	-0.003	-0.008***	-0.005
	(0.003)	(0.003)	(0.002)	(0.003)
Constant	1.095***	1.048***	1.201***	1.189***
	(0.044)	(0.049)	(0.047)	(0.051)
Observations	2880	2880	3360	3360
R^2	0.239	0.228	0.296	0.268

Standard errors in parentheses. 8 clusters in all columns.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.5: *Individual turnout decisions*. Standard errors are clustered at the session level.

erogeneity is estimated, as opposed to being imputed by assumption to each individual. Table B.6 reports the results of estimating our turnout regression via a probit model with random effects.

Once again all signs remain unchanged, but in two cases the modelling procedure affects the statistical significance of the parameter of the interaction term, increasing it if $m = 2, K = 4$, decreasing it if $m = 3, K = 2$. [6]

B.3.2 Minority victories

Between-subjects design: first treatments only

Table B.7 replicates the regression in Table 3.5 using only data from treatments run first in a session. The positive impact of CV on the share of seats won by the minority is confirmed in all parametrizations.

[6]The numbers reported in the table are raw coefficients from the probit regression.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
Turnout				
Minority	-0.886***	-0.219	-0.519**	-0.831***
	(0.258)	(0.220)	(0.224)	(0.225)
CV	0.159	-0.011	-0.361*	0.018
	(0.147)	(0.149)	(0.187)	(0.160)
CV × Minority	0.675**	0.931***	0.394	0.296
	(0.281)	(0.280)	(0.353)	(0.313)
Voting Cost	-0.033***	-0.034***	-0.040***	-0.039***
	(0.003)	(0.003)	(0.003)	(0.003)
CVfirst	-0.283	-0.143	0.140	-0.002
	(0.244)	(0.245)	(0.218)	(0.234)
Round	-0.040***	-0.012	-0.037***	-0.041***
	(0.009)	(0.009)	(0.008)	(0.009)
Constant	2.702***	2.522***	3.246***	3.502***
	(0.255)	(0.278)	(0.226)	(0.259)
$\ln(\sigma^2)$	0.253	0.256	0.175	0.317
	(0.231)	(0.215)	(0.214)	(0.207)
Observations	2880	2880	3360	3360

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.6: *Individual turnout decisions*. Estimated probit model with random effects.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
CV	0.190***	0.163***	0.218***	0.307***
	(0.037)	(0.044)	(0.054)	(0.050)
Round	-0.005	0.003	-0.002	-0.005
	(0.006)	(0.003)	(0.005)	(0.006)
Constant	0.088	0.139***	0.200***	0.196***
	(0.064)	(0.052)	(0.069)	(0.069)
Observations	411	413	467	466
R^2	0.083	0.120	0.097	0.168

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.7: *Share of seats won by the minority*. Data restricted to treatments run first in each session.

	(1) m2K2	(2) m2K4	(3) m3K2	(4) m3K4
CV	0.241**	0.202***	0.230***	0.224***
	(0.088)	(0.057)	(0.057)	(0.020)
CVfirst	0.045	-0.003	0.069	0.059
	(0.049)	(0.061)	(0.082)	(0.070)
CV × CVfirst	-0.164	-0.005	-0.068	-0.011
	(0.136)	(0.072)	(0.088)	(0.054)
Round	-0.001	-0.000	-0.002	0.001
	(0.003)	(0.002)	(0.002)	(0.002)
Constant	0.094*	0.161***	0.188**	0.098*
	(0.048)	(0.043)	(0.056)	(0.051)
Observations	1649	1641	1870	1861
R^2	0.094	0.165	0.082	0.131

Standard errors in parentheses. 8 clusters in all columns.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table B.8: *Share of seats won by the minority*. Standard errors are clustered at the session level.

Clustering

Table B.8 replicates the regression in Table 3.5, with standard errors clustered at the session level.

B.3.3 Individual tournout decisions: monotonicity violations

As reported in the text, we calculated for each participant the number of monotonicity violations: the minimum number of decisions that need to be modified for that participant's turnout to be fully monotonic in the voting cost realization. Because each participant plays 15 rounds of each treatment, the maximum possible number of violations to monotonicity is 7. Figure B.1 reports the CDF's of the number of violations for each parametrization, for both CV (the darker line) and MP (the lighter line). As comparison, we also report in the dashed line the CDF obtained from 10,000 simulations where the turnout decision is random and independent of the realized voting cost.

Under either voting rule, the CDF of monotonicity violations is strongly to the left of the CDF corresponding to random behavior. As stated in the text and most important for our purposes, the CDF's for MP and CV are very close and show no dominance relation. Kolmogoroff-Smirnov

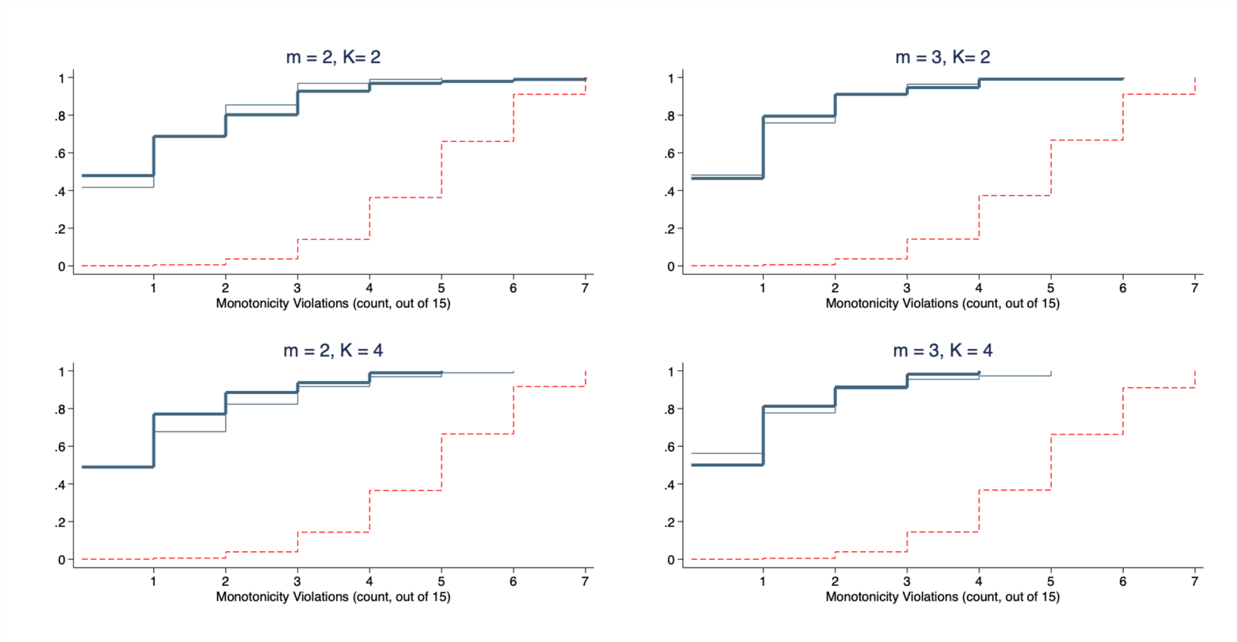


Figure B.1: *CDF's of monotonicity violations.* The darker line corresponds to CV; the lighter line to MP, and the dashed line to a simulation with random behavior.

(K-S) two-sample tests, corrected for discreteness, cannot reject the hypothesis of a common population for MP and CV, with p-values equal to 0.65 ($m2K2$), 0.33 ($m2K4$), 0.94 ($m3K2$), 0.58 ($m3K4$). The figure reports monotonicity violations aggregating over the two parties. The conclusion is unchanged if the data are organized by party.

Appendix C: Experimental Instructions for Chapter 1

C.1 Experimental Instructions for the *NumberPartial* Treatment

Thank you for participating in this experiment.

Please read and sign the consent form in front of you. Raise your hand if you have any questions.

In this experiment, you will make decisions in groups.

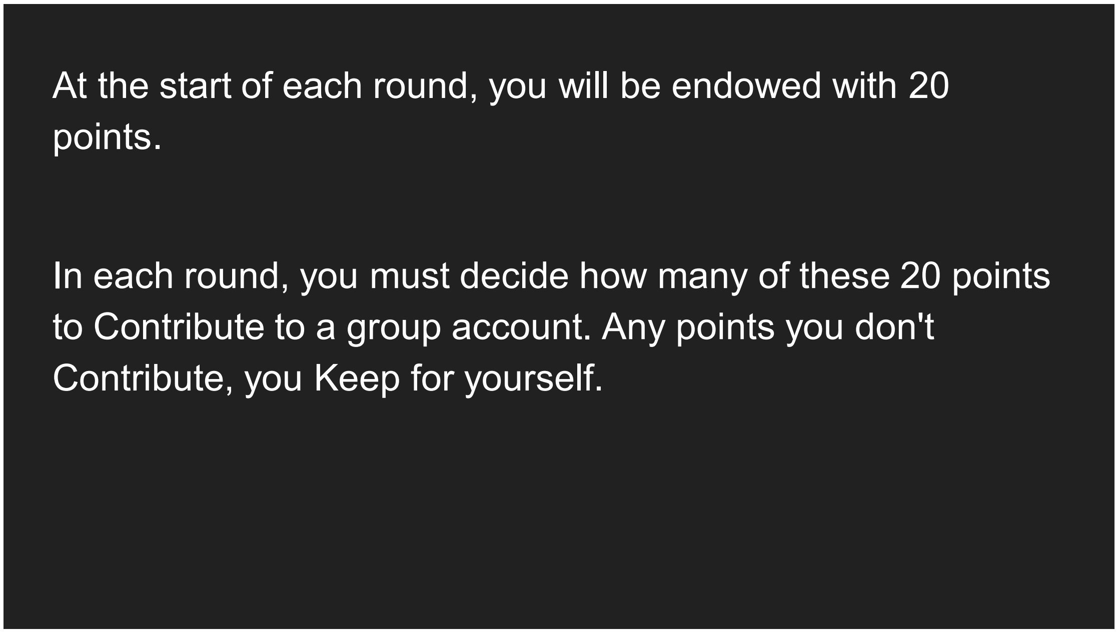
You will earn points throughout the course of the experiment.
How many points you earn depends both on your decisions and your group members' decisions.

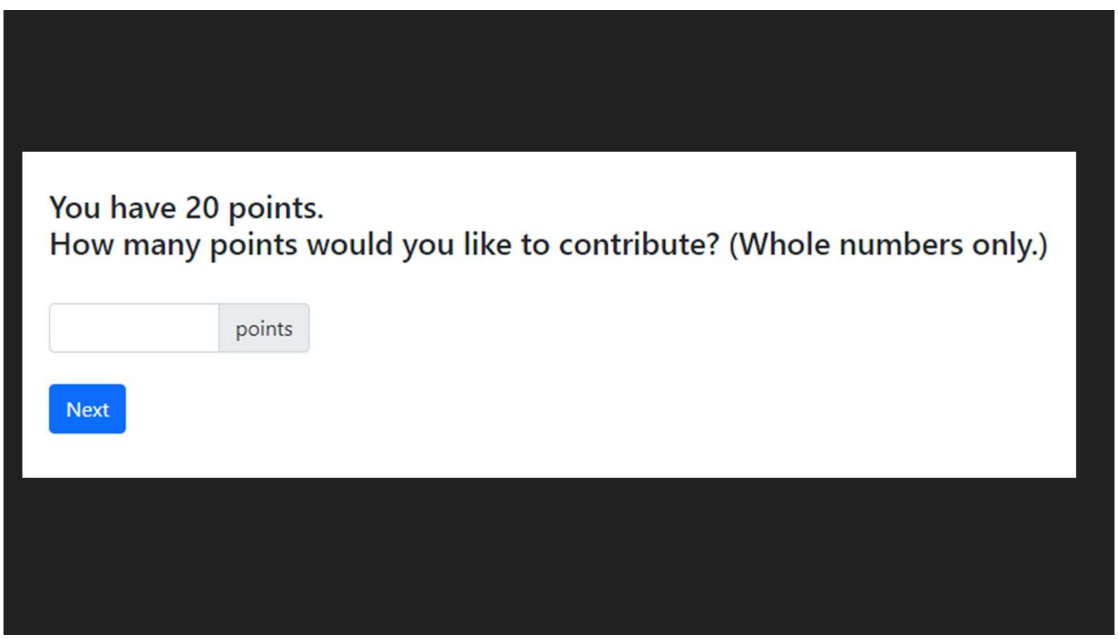
For being here, you are guaranteed to earn at least $5.

Your exact monetary earnings will depend on how many points you earn during the experiment.

At the start of this experiment, you will be randomly assigned to a group of 4 people (including yourself).

You will stay in this group for the entire experiment.


At the start of each round, you will be endowed with 20 points.
In each round, you must decide how many of these 20 points to Contribute to a group account. Any points you don't Contribute, you Keep for yourself.


You have 20 points.
How many points would you like to contribute? (Whole numbers only.)
points
Next

All contributions to the group account are **doubled**.

For example: if you Contribute 1 point, it becomes 2 points in the group account.

At the end of each round, the group account is evenly split among all four members of the group.

The number of points you earn each round is equal to:

1. How many points you Keep

   PLUS

2. Your Share of the Group Account

In other words, the number of points you earn in a round equals:

(20 - # points you Contribute)

PLUS

¼ * 2 * (# points you Contribute

+ # points your group members Contribute)

If you, and everyone else in your group, chooses to **Contribute 10 points**:

You Keep: **10** points

Your Share of Group Contribution:

¼ * 2 * (10+10+10+10) = **20** points

You Earn: 10 + 20 = <u>**30**</u> points

If you, and everyone else in your group, chooses to **Contribute 0 points**:

You Keep: **20** points

Your Share of Group Contribution:

¼ * 2 * (0+0+0+0) = **0** points

You Earn: 20 + 0 = <u>**20**</u> points

If everyone else in your group chooses to **Contribute 10 points** and you choose to **Contribute 12 points**:

You Keep: **8** points

Your Share of Group Contribution:

¼ * 2 * (12+10+10+10) = **21** points

You Earn: 8 + 21 = <u>**29**</u> points

If everyone else in your group chooses to **Contribute 10 points** and you choose to **Contribute 12 points**:

You Keep: **8** points

Your Share of Group Contribution:

¼ * 2 * (12+10+10+10) = **21** points

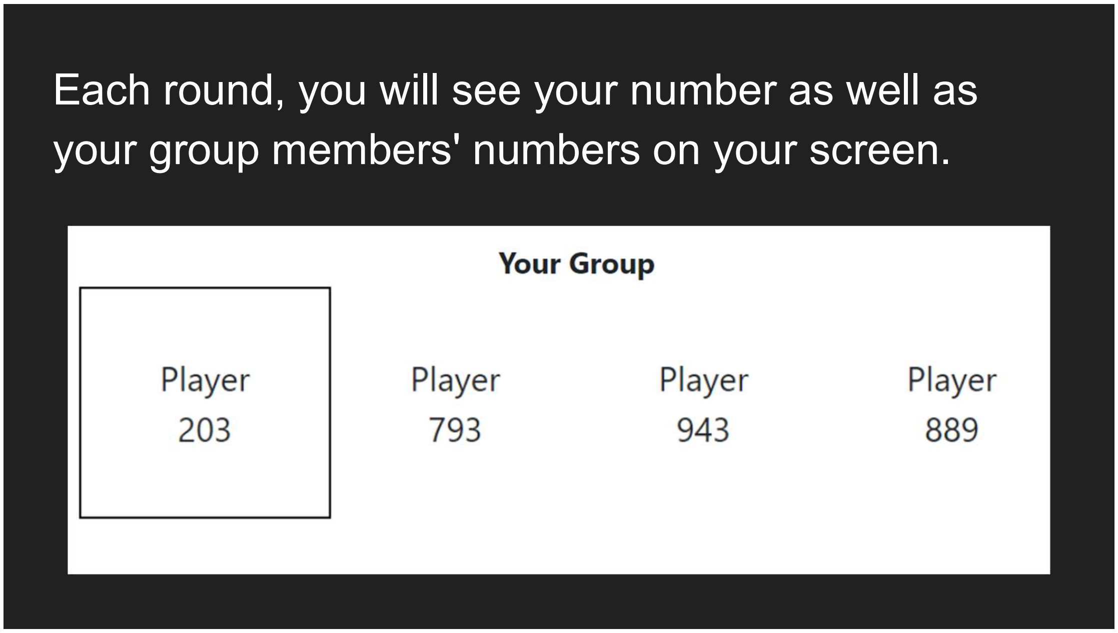
You Earn: 8 + 21 = **<u>29</u>** points

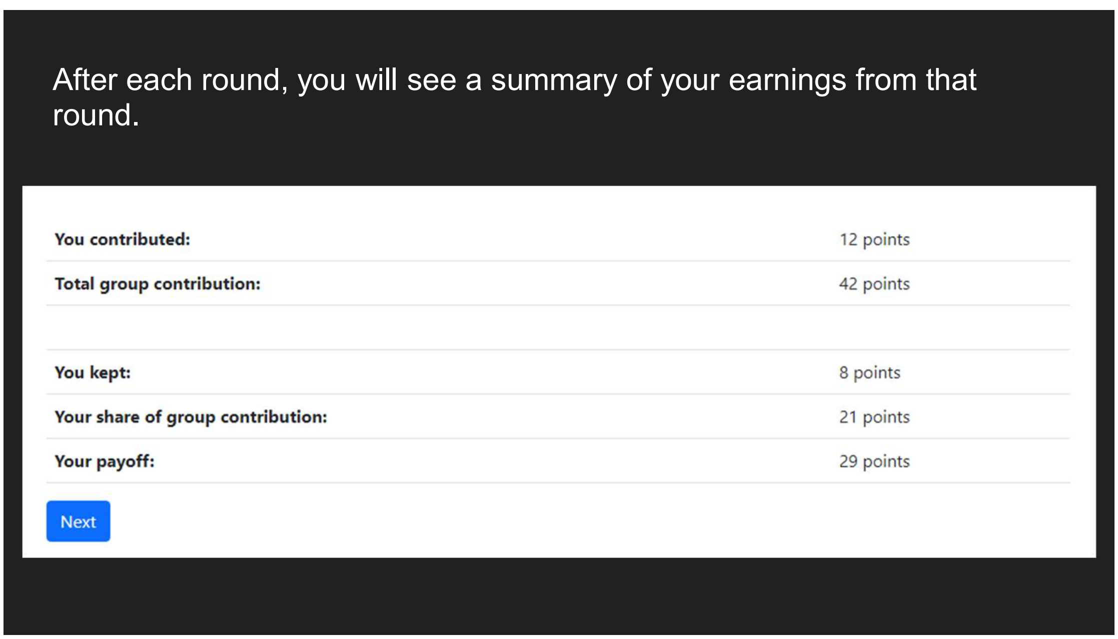
(Your Group Members Earn: 10 + 21 = 31 points)

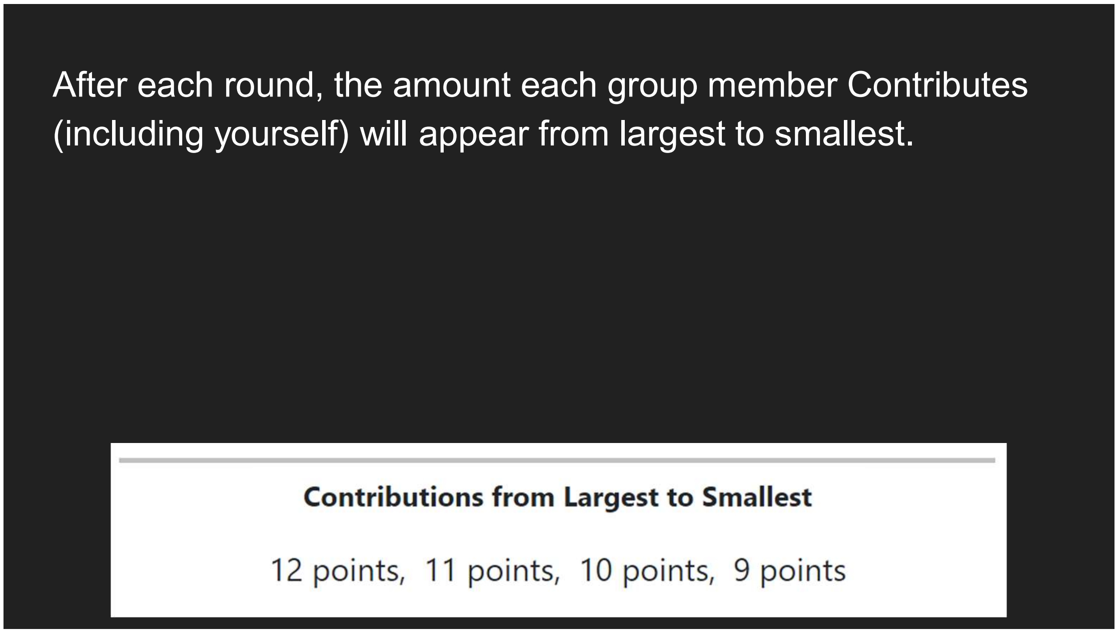
At the start of the experiment, every player will be randomly assigned a three-digit number.

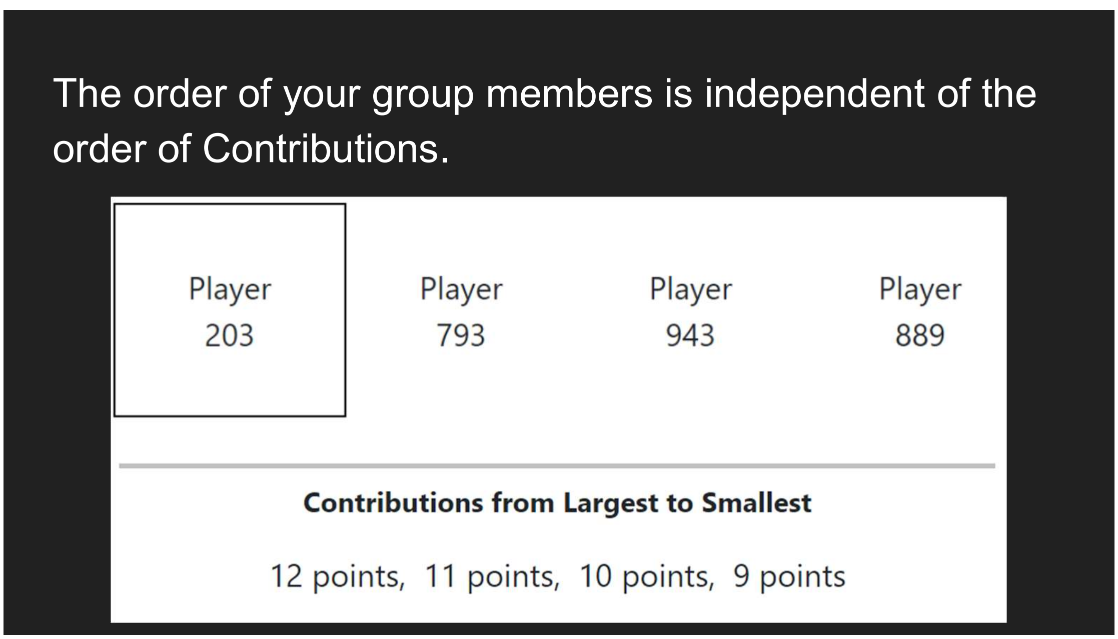
No two players will have the same number.

Only you will know what your number is.


Each round, you will see your number as well as your group members' numbers on your screen.
Your Group
Player 203
Player 793
Player 943
Player 889


After each round, you will see a summary of your earnings from that round.
You contributed: 12 points
Total group contribution: 42 points
You kept: 8 points
Your share of group contribution: 21 points
Your payoff: 29 points
Next

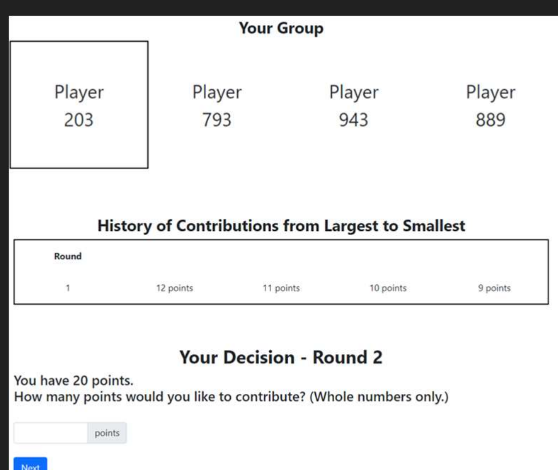

After each round, the amount each group member Contributes (including yourself) will appear from largest to smallest.
Contributions from Largest to Smallest
12 points, 11 points, 10 points, 9 points


The order of your group members is independent of the order of Contributions.
Player 203
Player 793
Player 943
Player 889
Contributions from Largest to Smallest
12 points, 11 points, 10 points, 9 points

You are not obligated to tell anybody what your number is, even after the experiment is over.

Nobody else, not even the experimenter, will be able to link your number back to you.

In later rounds, the computer will remind you of your group's Contributions in order from largest to smallest, from all previous rounds.



There will be **<u>32 rounds</u>** in total in this experiment, all of which proceed as previously described.

Your monetary earnings at the end of the experiment depend on the total number of points you earn in all 32 rounds.

<u>Every 50 points you earn equals $1</u>. (1 point = 2 cents)

<u>SUMMARY</u>

Same group of 4 people

Endowment: 20 points per round

Contributions are **<u>doubled</u>** in the group account

32 rounds in total

50 points = 1 dollar

Appendix D: Experimental Instructions for Chapter 2

This section reproduces the text of experimental instructions.

D.1 Welcome

Thank you for participating in this study. This study will take approximately **10 minutes** to complete. Please read all instructions carefully. You will receive a participation fee of $1 when you complete the study. In addition, there is a 20% chance you will be selected to receive a bonus payment. The bonus payment will depend on your choices in the study, and the specifics will be described in detail on a subsequent page.

D.2 Dictator Game - Receivers

Description:

This scenario is about allocating a sum of money between yourself and one other person who is recruited on MTurk the same way as you. The two of you will play the role of Person A and Person B, and these roles will be randomly assigned.

Person A starts with **$10**, while Person B starts with **$0**. Person A can decide how much of this money to keep for themself and transfer to Person B.

Person A receives the amount that they decide to keep while Person B receives the amount passed on to them by Person A. Person B does not get to make any decision and simply receives the amount passed on to them.

Your role:

You are assigned to be **Person B**, and another MTurk participant will be Person A.

[*If Control Receiver in Experiment R:* When Person A makes their money allocation decision, they will see the **avatar** that you designed to represent yourself in this study.]
[*If Treatment Receiver in Experiment R:* On the next page, you will design an **avatar** to represent yourself in this study. **Person A will be shown this avatar when they make the money allocation decision.**]

Bonus Payment:

Recall that there is a 20% chance that you will be randomly selected to receive a bonus payment. If you are selected to receive the bonus payment, your bonus payment will be the amount that Person A passes on to you.

Regardless of whether you are chosen to receive a bonus payment, you will receive a participation fee of $1 when you complete the study.

D.3 Dictator Game - Senders

Description:

This scenario is about allocating a sum of money between yourself and one other person who is recruited on MTurk the same way as you. The two of you will play the role of Person A and Person B, and these roles will be randomly assigned.

Person A starts with **$10**, while Person B starts with **$0**. Person A can decide how much of this money to keep for themself and transfer to Person B.

Person A receives the amount that they decide to keep while Person B receives the amount passed on to them by Person A. Person B does not get to make any decision and simply receives the amount passed on to them.

Your role:

You are assigned to be **Person A**, and another MTurk participant will be Person B.

[*If Sender in Experiment R:* When you make your money allocation decision, you will see an **avatar** that Person B created to represent themself in the study. You will take part in **3** rounds of the above scenario. In each round, you will be matched with a new Person B and see their avatar when making your money allocation decision.]

Bonus Payment:

Recall that there is a 20% chance that you will be randomly selected to receive a bonus payment. **If you are selected to receive the bonus payment, your bonus payment will be the amount that you decide to keep for yourself. Additionally, the other participant (Person B) will receive the amount passed on by you.**

Regardless of whether you are chosen to receive a bonus payment, you will receive a participation fee of $1 when you complete the study.

D.4 Avatar Rating Task

In this part of the experiment, you will be asked to rate 15 cartoon avatars on different dimensions. Each dimension will be characterized by a **pair of words**.

For each word pair, please find the word that best fits the avatar. Additionally, please pick how well you think that word fits the avatar.

Note: the order of the words can change between avatars!

D.5 Avatar Creation Task

Use the dropdown menus to design the avatar that will represent you in the study.

The avatar at the top of the page will auto-update based on the choices from the dropdown menus.

You must make at least one choice from one dropdown menu.

When you are finished designing your avatar, click the Next button at the bottom of the page. The button will appear automatically after **150 seconds**.

Appendix E: Experimental Instructions for Chapter 3

E.1 Experimental Instructions for a session with 12 subjects and $m = 2$

In this experiment, you are an eligible voter in an election called to fill several seats.

You will be assigned to one of two parties, either the **Purple** party or the **Orange** party.

If your party wins positions, you win points. However, voting is costly, so if you vote, you lose points.

Your decision in each round will be whether to **vote** or **not vote**.

The experiment will have 4 *Parts*, each with 15 *rounds.*

At the end of each Part, 1 round will be randomly drawn, and the points you earned in those rounds will be converted into monetary earnings.

The conversion rate from points to dollars is: 100 points = $1.50 dollar.

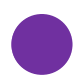
Your final earnings will also include the show-up fee of $5.

Part 1

Part 1 consists of 15 rounds.
Each round is a new election and will develop as follows:

At the beginning of Part 1, you will be randomly assigned to a party: either **Orange** or **Purple**.

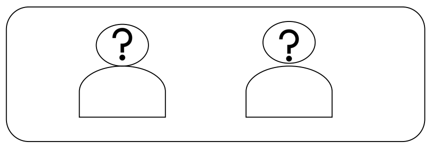
or



You will be part of a group of **6 voters**: **2 Purple** voters and **4 Orange** voters.

Your group will vote to elect a commission with 2 seats.

2 seats

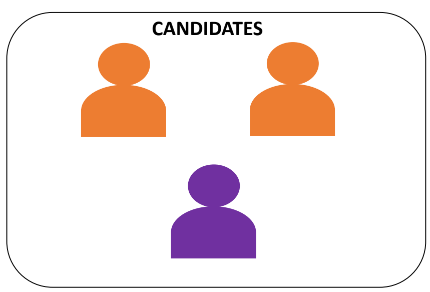


The 2 candidates with most votes win. If there is a tie, it is broken randomly by the computer.

You win **400 points** if your party wins **both** positions, **200 points** if it wins **1** position, and **0** if it wins **none**.

To fill these seats, the **Orange** party has nominated **2 candidates**.

The **Purple** party has nominated **1 single candidate.**



The computer screen will remind you of this information, and will look as follows.

Your party affiliation is ORANGE.

This is Round 1.

Your group consists of 4 Orange members and 2 Purple members.

You are asked to elect a committee of size 2. You have 2 votes.

There are 2 Orange candidates and 1 Purple candidates.

You win 400 points if your party wins both positions; 200 if your party wins 1 position;
and 0 if all positions are won by the other party.

Continue

Whether you are **Orange** or **Purple**, your decision is whether to **vote** or **not vote**.

You are given an **endowment of 100 points**. But if you vote you suffer a voting cost.

The cost is random and can be any **integer between 0 and 100 with equal probability.**

Your cost is **independent** of everyone else's cost. Thus, your cost gives you no information about their costs.

You will be shown your cost before you choose to **vote** or **not vote**.

Your party affiliation is ORANGE.

This is Round 1.

Your voting cost for this Round only is

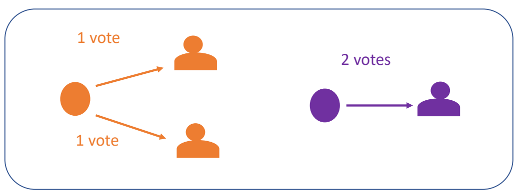
22

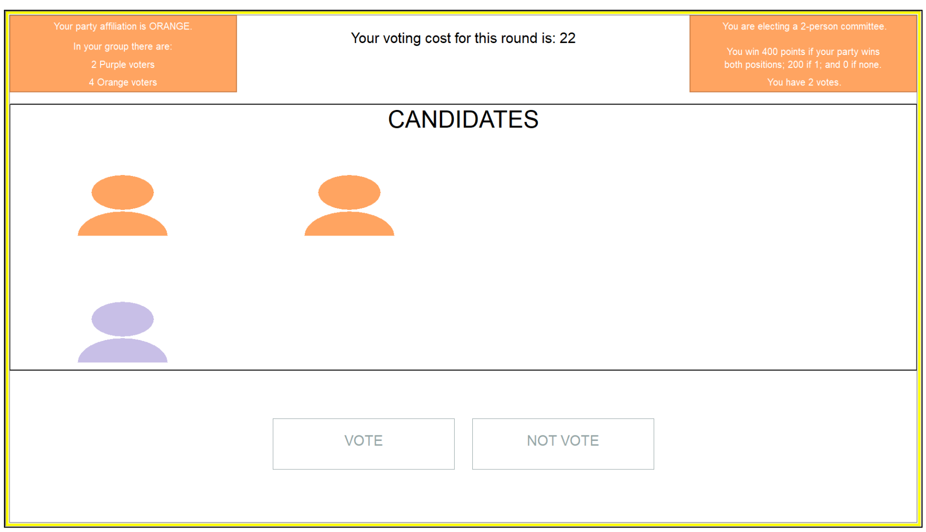
Because 2 positions are open, **each voter has 2 votes**.

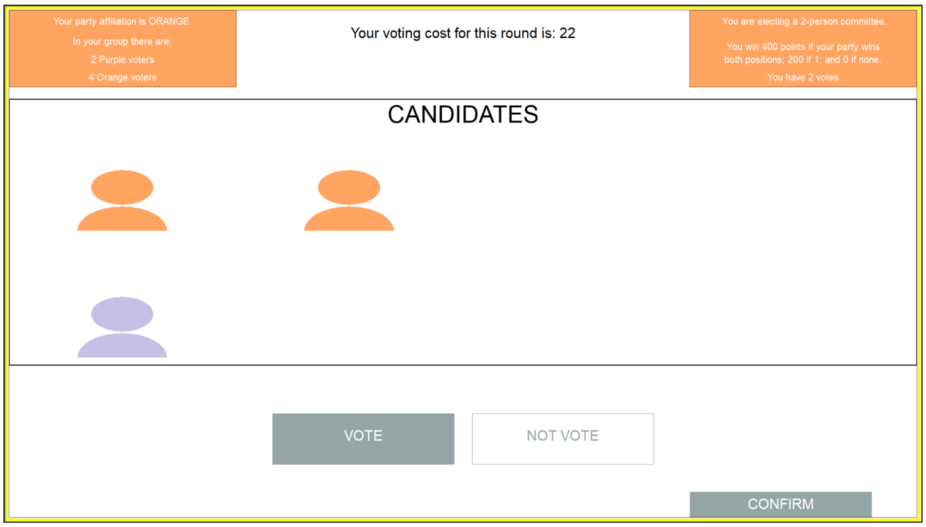
When you vote, all your votes are shared equally by all candidates of your party.

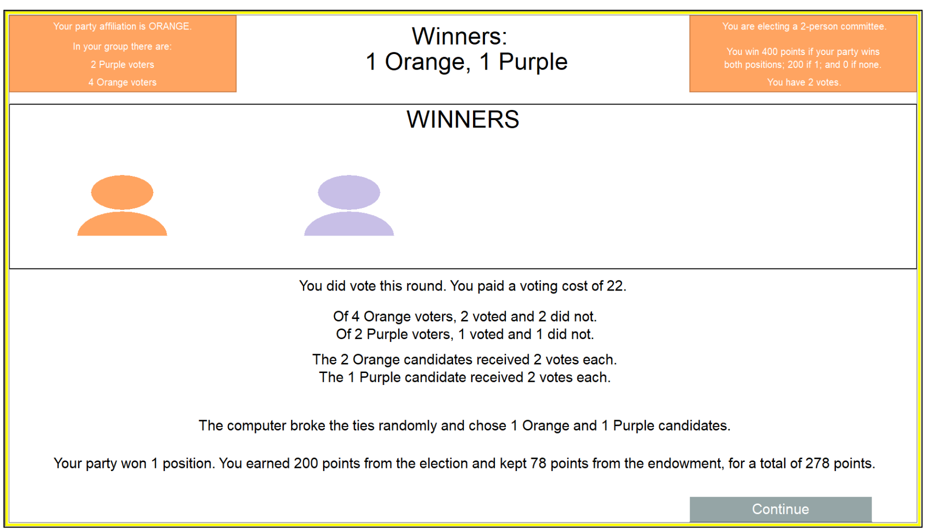
If your party is **Orange** and you vote, you cast **1 vote** for each of the 2 **Orange** candidates.

If your party is **Purple** and you vote, you cast **both of your votes** for the single **Purple** candidate.






Your party affiliation is ORANGE.
In your group there are:
2 Purple voters
4 Orange voters
Your voting cost for this round is: 22
You are electing a 2-person committee.
You win 400 points if your party wins both positions; 200 if 1; and 0 if none.
You have 2 votes.
CANDIDATES
VOTE
NOT VOTE
CONFIRM

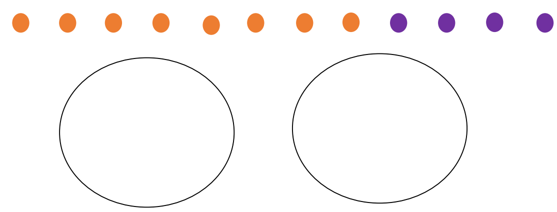

Your party affiliation is ORANGE.
In your group there are:
2 Purple voters
4 Orange voters
Winners:
1 Orange, 1 Purple
You are electing a 2-person committee.
You win 400 points if your party wins both positions; 200 if 1; and 0 if none.
You have 2 votes.
WINNERS
You did vote this round. You paid a voting cost of 22.
Of 4 Orange voters, 2 voted and 2 did not.
Of 2 Purple voters, 1 voted and 1 did not.
The 2 Orange candidates received 2 votes each.
The 1 Purple candidate received 2 votes each.
The computer broke the ties randomly and chose 1 Orange and 1 Purple candidates.
Your party won 1 position. You earned 200 points from the election and kept 78 points from the endowment, for a total of 278 points.
Continue

After the practice rounds, Part 1 will consist of 15 rounds.

Each round is a new election:

1. You will maintain the same party affiliation (**Orange** or **Purple**) for all of Part 1.

2. At each round, you will be rematched randomly into a group of 6 voters: **4 Orange** and **2 Purple**. Because groups are rematched, you will **not** be with the same five other subjects.



3. The **groups are independent**. What happens in one group does not affect the others.

4. The election will proceed as before (2 open seats, 2 votes per voter, **2 Orange candidates**, **1 Purple candidate**)

5. All random **voting costs will be drawn again**. Each voter will receive a new endowment of 100 points.

6. You and everyone else will decide again whether to vote or not.

SUMMARY

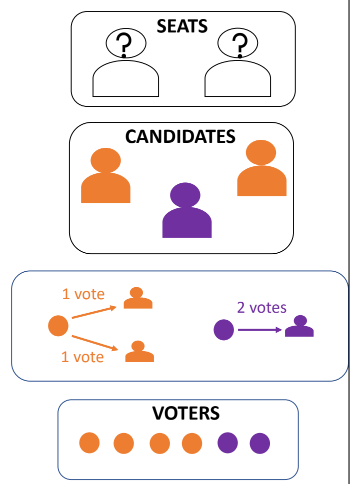
Your group votes to fill **2 open seats** and **each voter has 2 votes**.

You win 400 points if your party wins both seats, 200 if 1, 0 if none.

There are **2 Orange** candidates and **1 Purple** candidate.

Because each voter has 2 votes, an **Orange voter** who votes casts **1 vote** for each Orange candidate; a **Purple voter** who votes casts **2 votes** for the single Purple candidate.

The group consists of **4 Orange voters** and **2 Purple voters.**



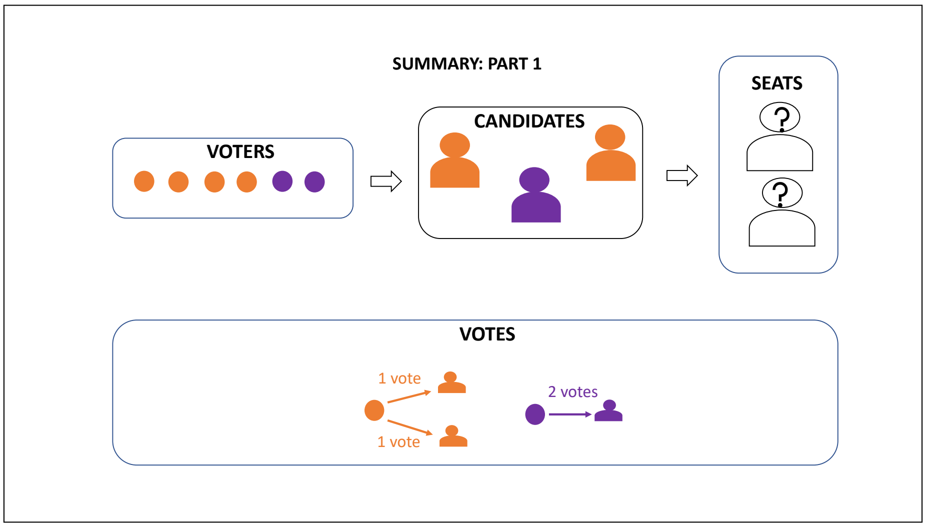
Voting is costly.

If you vote, you pay your voting cost whether or not your party wins any position.

If your party wins any position, you will receive points whether you voted or not.

Your party is more likely to win positions if you vote.

Your decision will be whether to vote or not.


SUMMARY: PART 1
VOTERS
CANDIDATES
SEATS
?
?
VOTES
1 vote
1 vote
2 votes

Part 2

Part 2 is very similar to Part 1.

The difference is in the number of candidates:

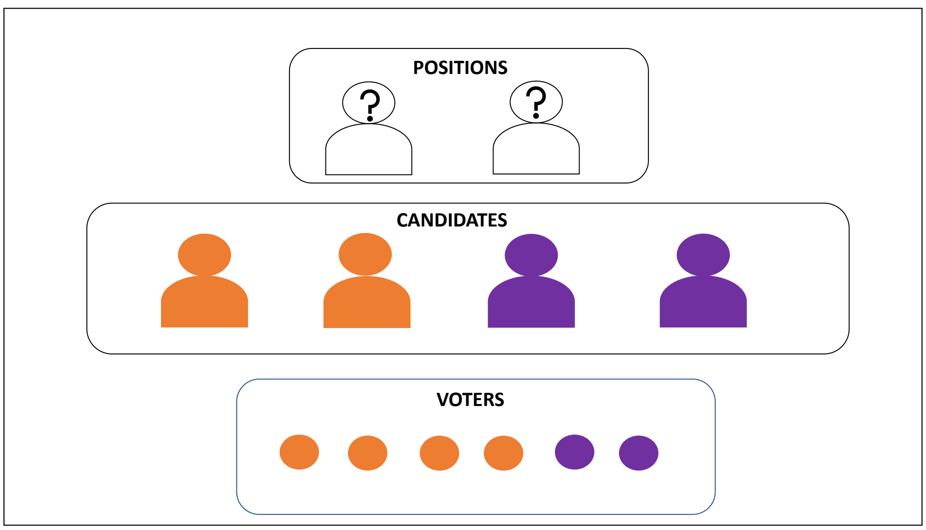
In part 2, both parties have nominated **2 candidates** each.

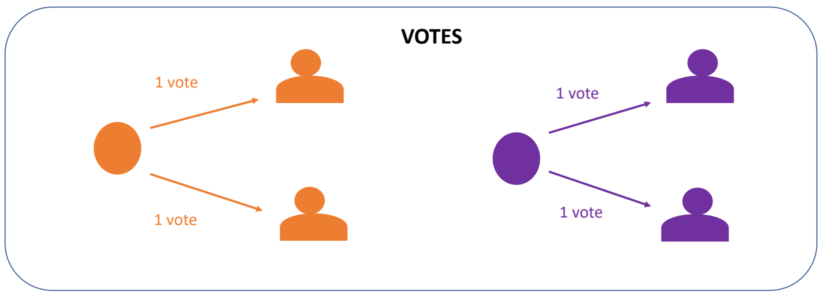
If you vote, **you cast 1 vote for each of your party's candidates**.

The two winning candidates will be the two candidates with most votes, as before.

You will be randomly reassigned a party affiliation, either **Purple** or **Orange**, which may differ from Part 1 and that you will keep for the whole of Part 2.

Everything else remains identical.



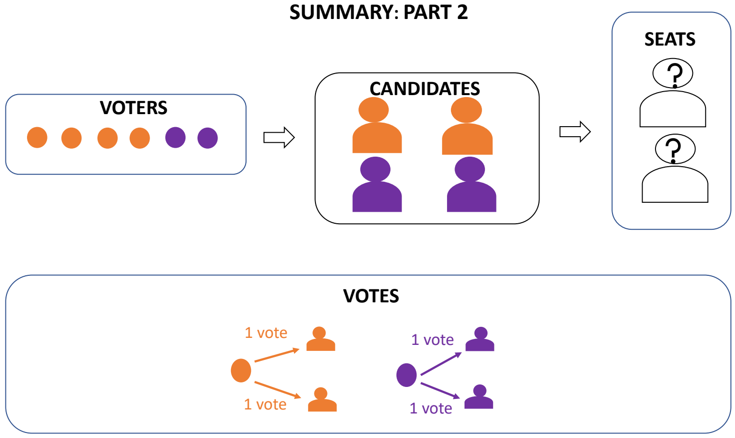


Part 2 will last 15 rounds.

Remember that:

- groups are reshuffled at each round.
- voting costs are redrawn at each round.

SUMMARY: PART 2



Part 3

Part 3 is very similar to Part 2.

There are three differences:

1. In Part 3, your group will elect a **4-person commission:** the election is for 4 open seats.

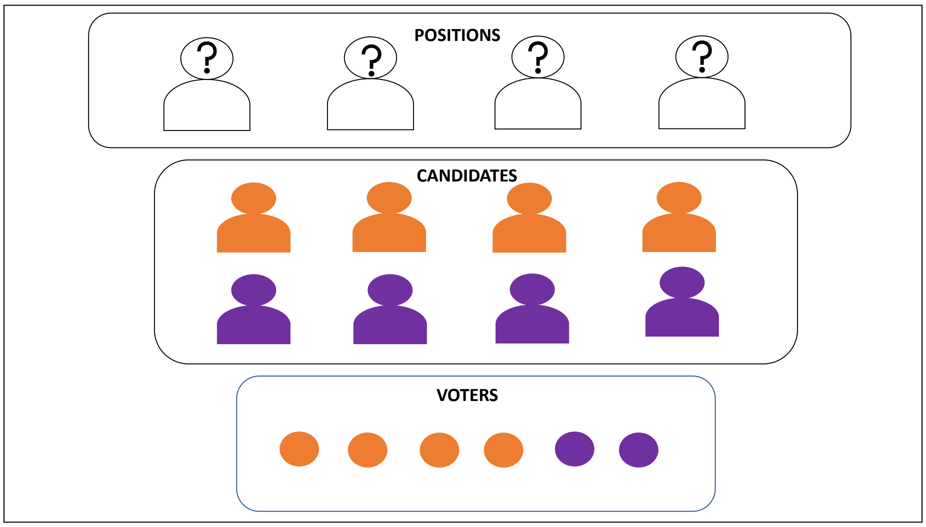
2. Each voter has **4 votes**.

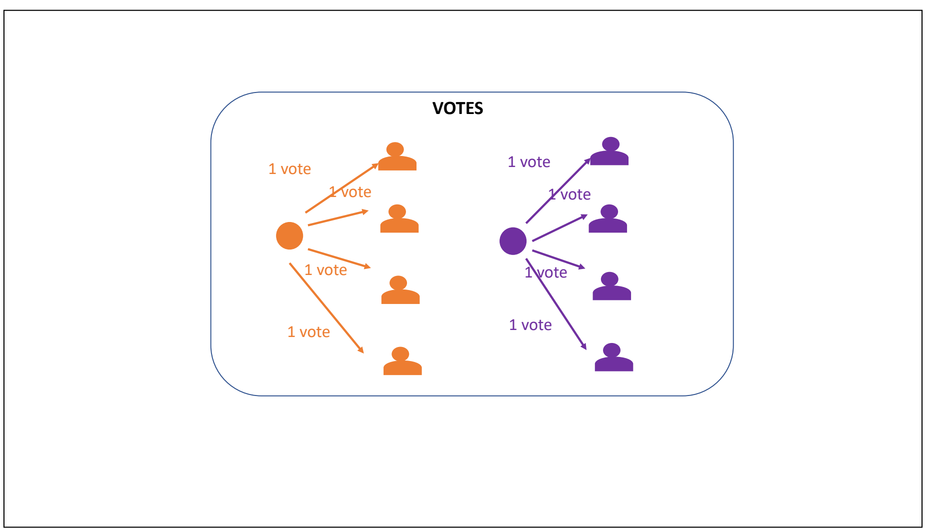
3. Each party has nominated **4 candidates.**

If you vote, you cast **1 vote for each of your party's candidates**.

The 4 winning candidates will be the 4 candidates with most votes.

In addition to the 100 points round endowment, you win 400 points if your party wins 4 positions, 300 if it wins 3, 200 if 2, 100 if 1, and 0 if it wins none.


POSITIONS
CANDIDATES
VOTERS

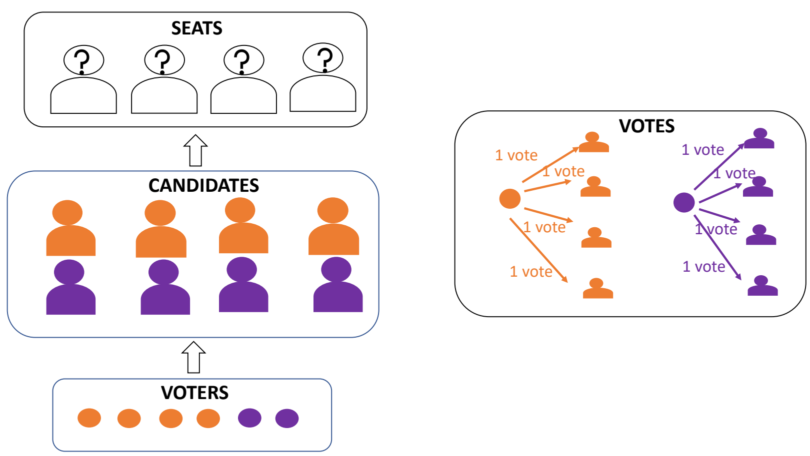

VOTES
1 vote
1 vote
1 vote
1 vote
1 vote
1 vote
1 vote
1 vote

You will be randomly reassigned a party affiliation, either **Purple** or **Orange**, which may differ from Part 2 and that you will keep for the whole of Part 3.

Everything else remains identical.

Part 3 will last 15 rounds.

SUMMARY: PART 3



Part 4

Part 4 is very similar to Part 3.

The difference is in the number of candidates:

The **Orange** party has nominated **3 candidates**.

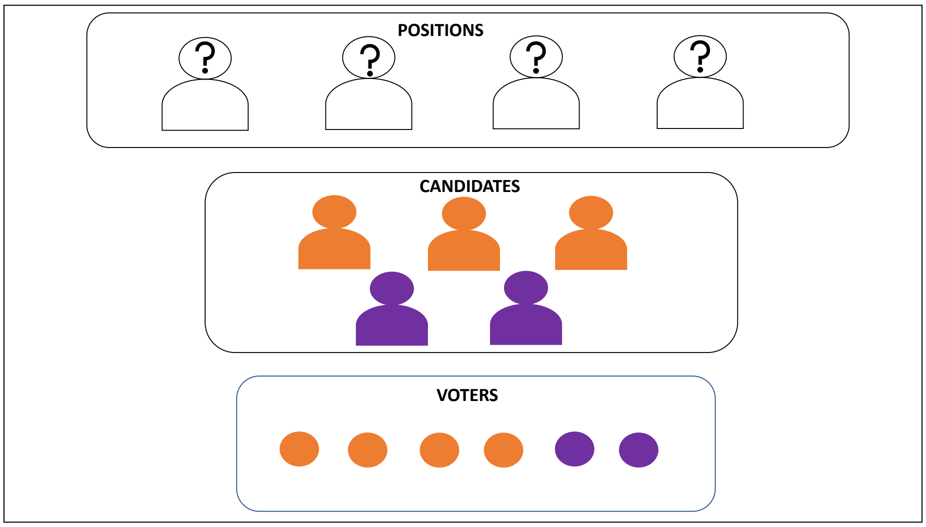
The **Purple** party has nominated **2 candidates.**

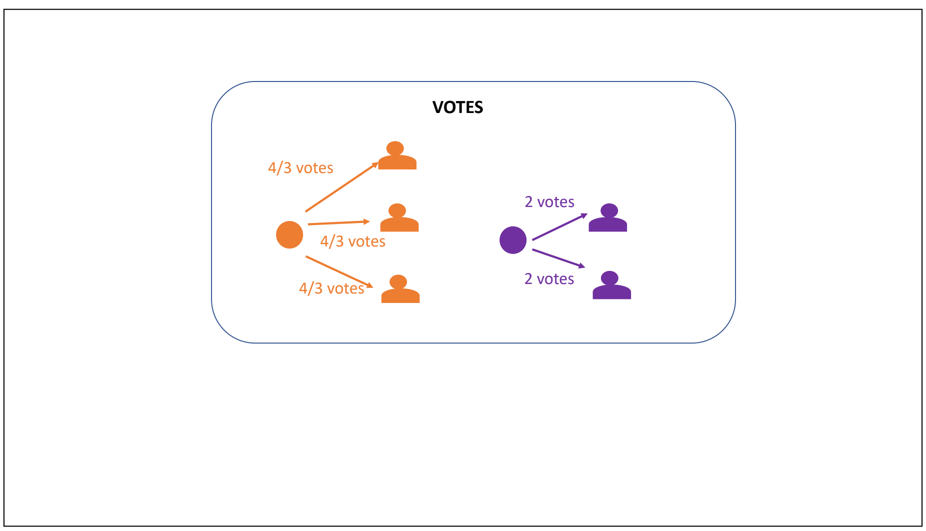
Because every voter has 4 votes:

If you vote and are **Orange**, you cast **4/3 votes** for each **Orange candidate**.

If you vote and are **Purple**, you cast **2 votes** for each **Purple candidate**.

The 4 winning candidates will be the 4 candidates with most votes, as before.


POSITIONS
CANDIDATES
VOTERS


VOTES
4/3 votes
4/3 votes
4/3 votes
2 votes
2 votes

You will be randomly reassigned a party affiliation, either **Purple** or **Orange**, which may differ from Part 3 and that you will keep for the whole of Part 4.

Everything else remains identical.

Part 4 will last 15 rounds.