PREFERENCES FOR POWER

ELENA PIKULINA AND CHLOE TERGIMAN

ABSTRACT. In many social contexts, the ability to decide or influence someone else's outcomes is associated with benefits for the decision-maker, for example higher compensation for managers or public recognition for politicians. Using a novel game, we demonstrate that individuals enjoy power – the ability to determine outcomes for others – by and of itself. In fact, they are willing to accept a lower payoff for themselves in exchange for power, even in the absence of any positive externalities. We show that individuals' preferences for power are different from, and cannot be explained by, their social preferences.

Date: **First Version: May 2017. Current version: October 2017**.

Key words and phrases. Preferences for power, private benefits of control, social preferences, other-regarding preferences, laboratory experiment

JEL: C91, D01, D03, M21.

Pikulina: elena.pikulina@sauder.ubc.ca, Sauder School of Business, University of British Columbia. Tergiman: cjt16@psu.edu, Smeal College of Business, Penn State University.

We thank Nageeb Ali, Gary Charness, Emel Feliz-Ozbay, Zack Grossman, Yoram Halevy, Paul J. Healy, Holger Herz, Tony Kwasnica, Yusufcan Masatlioglu, Marie-Claire Villeval, Ryan Oprea, Erkut Ozbay, Carolin Pflueger, Ran Schorrer, Ron Siegel, Philippe Tobler, Neslihan Uler, and Sevgi Yuksel for lively discussions and valuable feedback. We also benefited from comments from seminar participants at the Sauder School of Business (University of British Columbia), Simon Fraser University, Pittsburgh University, The Technion (Haifa, Israel), GATE-LSE (Lyon, France), University of Zürich, University of Maryland, The Pennsylvania State University as well as participants at the following conferences: the Southwest Experimental and Behavioral Economics Workshop (University of California Santa Barbara), the Bay Area Behavioral and Experimental Economics Workshop (Santa Clara University), ESA San Diego, the Tilburg Institute for Behavioral Economics Research Conference and the Center for Experimental Social Science 15th Anniversary Conference (NYU) . Pikulina and Tergiman are also very grateful for generous funding from the Social Sciences and Humanities Research Council of Canada (SSHRC).

> “When a moderate degree of comfort is assured, both individuals and communities will pursue power rather than wealth: they may seek wealth as a means to power, or they may forgo an increase of wealth in order to secure an increase of power, but in the former case as in the latter their fundamental motive is not economic.”
>
> — Bertrand Russell, Power

1. Introduction

Rational economic agents with standard preferences are interested in controlling the fates of others only as long as such power gives them material benefits, for example, increases their payoffs, expands their choice set, or decreases risk and uncertainty. Here, we eliminate any material benefits of power to the decision-maker and study its intrinsic value instead.

As it relates to the interaction between people, the Oxford Dictionary defines power as "the capacity or ability to direct or influence the behavior of others or the course of events."[1] In a principal-agent context, or one of manager-employee, two aspects of power are most prevalent. One is a manager’s ability to decide on an employee’s tasks and responsibilities. Another, which we study in this paper, is the manager’s ability to determine an employee’s pay, whether in the form of a bonus or a promotion that would directly affect the employee’s remuneration. Of course, power is not limited to principal-agents settings, and individuals, organizations, and states can have power over each other to varying degrees and in various contexts.[2] In this paper, we study power as the ability to determine someone else’s compensation.

Authority, power, control, and autonomy are notions that are often confounded. For example, when a manager has authority over employees, her authority includes *power*, the ability to determine their payoffs and responsibilities, *control*, the ability to determine her own payoff and responsibilities, and *autonomy*, the enjoyment of non-interference in her affairs by others.[3] Here, we focus solely on power and evaluate it independently of these other factors.

[1]See Oxford Dictionary https://en.oxforddictionaries.com/definition/power.

[2]In the United States, presidents can exercise immense power, especially in times of war and national emergency. On January 1, 1863, President Abraham Lincoln issued the Emancipation Proclamation, an order that freed 3 million enslaved people; on February 19, 1942, Franklin D. Roosevelt signed executive order 9066 that cleared the way for Japanese-Americans to be sent to concentration camps for the duration of the World War II. In times of piece however the power of executive orders over individuals’ legal status and mobility is more limited and is usually subject to judicial review and has to be supported by the Congress.

[3]Control and autonomy are not synonymous. Consider, for example, the case where the manager’s payoff is determined randomly. In this case, she has no control but does have autonomy.

The main challenge in isolating individuals' preferences for power – the ability to determine payoffs of others – and in estimating their willingness to pay for it comes from the fact that people may put a non-zero weight on those payoffs. In other words, they may have social preferences that are independent of their preferences for power. For example, someone may not particularly enjoy choosing the payoff of someone else, but may enjoy the resulting payoff distribution. We design an experiment to specifically address this challenge: our experimental design allows us to separate individuals' willingness to pay for power from their willingness to pay to implement their social preferences.

We introduce a new game, the "Power Game." In the Power Game, there are two types of players, A and B, who are matched in pairs. Only type A players make decisions, and these decisions determine the payoffs of both the A player and the B player she is matched with. The Power Game is in two Parts. In Part I, A has the choice between two options. In the first option, B receives a pre-specified amount E_A, and A can choose her own payoff in the $[0, E_A]$ interval. A rational player always chooses E_A for herself, hence, the resulting allocation is (E_A, E_A). Player A's second option is to receive $E_A - p$ and obtain the right to choose a specific payoff for player B in the $[0, E_B]$ interval. In other words, by paying p, A obtains the right to determine the payoff of B. If A pays, ($E_A - p$, x_B^*) is the resulting allocation, where x_B^* is what A chose for B in the $[0, E_B]$ interval. Because Part I has several rounds and p varies from round to round, we can determine individuals' willingness to pay for the right to determine B's payoff.

In Part II, player A makes choices between two payoff pairs that determine payoffs for herself and for player B. Some of these pairs are determined by her actions in Part I.[4] If in Part I a player chose to pay price p, then in the corresponding round of Part II she has to choose between (E_A, E_A) and $(E_A - p, x_B^*)$, where x_B^* is what she chose for B in Part I. In other words, she has the choice between (E_A, E_A), the allocation that a payoff-maximizing player would have chosen, and the allocation she actually chose in Part I, $(E_A - p, x_B^*)$. If in Part I she chose not to pay price p, then in the corresponding round of Part II she has to choose between (E_A, E_A) and $(E_A - p, E_A + 2p)$. In other words, she has the choice between the allocation she actually chose, (E_A, E_A), and a more efficient allocation, $(E_A - p, E_A + 2p)$, that she could have chosen.[5]

There are two key features in the Power Game that allow us to identify subjects with power preferences. First, while player A has power over B in both Parts of the Power Game, she faces different trade-offs between power and her own payoff in Parts I and

[4]In our experimental implementation, players also make choices between payoff pairs unrelated to their decisions in Part I.

[5]For players that do not pay p in Part I, there is no way to know what allocation they might prefer to (E_A, E_A), if any. For that reason, in Part II, we offer them an efficient allocation $(E_A - p, E_A + 2p)$ as an alternative to their actual choice, (E_A, E_A).

II. It is only in Part I that A can buy more power, because in Part II, A's power is fixed regardless of her choice to pay. In Part I, when A pays p, she obtains the right to choose B's payoff precisely, and can choose any payoff she pleases within the interval $[0, E_B]$. An individual with preferences for power should be able to recognize this explicit power-payoff trade-off: paying p increases A's power over B. In contrast, Part II of the Power Game does not offer such a trade-off. When A gives up p in Part II and chooses the payoff pair with the lower payoff for herself, it does not change her power over B but simply implies a different, fixed, payoff for B. The second key feature of our design is that all Part I rounds have a corresponding round in Part II that presents player A with her Part I allocation and an alternative she could have chosen. These two features allow us to determine why a player paid in Part I. Did she pay because she desired a specific distributional outcome $(x_A - p, x_B^*)$, i.e., has social preferences? Or did she pay because she enjoyed the power of choosing B's payoff in $[0, E_B]$ but in fact attached little importance to her actual choice of x_B^*?

While players with standard preferences, i.e., payoff-maximizing players, never pay in Part I or in Part II of the Power Game, players who value power or have social preferences pay non-zero prices in Part I. It is these players' paying behavior in Part II that allows us to distinguish between these two preference classes. If A's choices in Part I are the result of her social preferences and she does not place any value on the process by which final allocations are attained, i.e., does not place any value on her ability to choose a specific payoff for B, then in Part II she should still prefer $(E_A - p, x_B^*)$. In other words, player A should be willing to pay price p to implement her social preferences irrespective of whether she picks B's payoff herself from the $[0, E_B]$ interval as in Part I, or whether the exact same payoff is exogenously given as in Part II. If, in contrast, in Part I, player A pays only to increase her power over B, then in Part II she should prefer (E_A, E_A), since paying in Part II does not lead to any additional power but simply lowers A's payoff. Thus, if a player reverses her choices in Part II and chooses (E_A, E_A) instead of $(E_A - p, x_B^*)$, the allocation she implemented in Part I, then she must have preferences for power. By comparing how much subjects are willing to pay in Parts I and II of the Power Game, we are able to identify their preferences for power and/or their social preferences.

Our main finding is that a large fraction of our subjects, specifically 42 percent, have preferences for power. In Part I, these subjects are willing to pay over 10 percent of their potential payoff to be able to choose payoffs for B, but they are willing to pay nothing to implement the *same* allocations in Part II, when additional power is not attainable. We also find that about 24 percent of our subjects have standard preferences, i.e., they do not attach any value to power or payoffs of others and never pay in either Part of the Power

Game. Power and social preferences are not mutually exclusive and 14 percent of our subjects have both. They have positive but different willingnesses to pay in Parts I and II of the Power Game. Among these subjects, the majority enjoy power over others while the rest are power-averse. Only 9 percent of our subjects have social preferences and are indifferent towards power, since they have identical willingnesses to pay in Parts I and II of the Power Game. We therefore show that a substantial majority of our subjects value power beyond its instrumental worth.[6]

We provide evidence that our Power-Game-based preference classification indeed captures differences in preferences across subjects. Since our classification depends only on the difference in subjects' willingnesses to pay across Parts I and II of the Power Game, we can use it to predict subjects' choices in other dimensions. We show that subjects we have classified as having social preferences, regardless of their attitude towards power, are consistent in the amounts they give to type B players. In contrast, subjects with power preferences and no social preferences exhibit much more variation in their giving behavior both within and across subjects. In addition, we show that these classes also predict subjects' decisions in tasks that are unrelated to Part I of the Power Game. More specifically, in the absence of power, subjects with power preferences and no social preferences behave much like subjects with standard preferences, i.e., they maximize their own payoff, while those with social preferences do not.

Our experimental results are closely related to the recent experimental literature on individual preferences for control and decision rights. Owens, Grossman and Fackler (2014) find that when asked whether to bet on their own performance or on their partners' performance in a quiz, people prefer to bet on themselves. Moreover, they are willing to sacrifice up to 15 percent of their expected payoff to retain control over their own outcome rather than delegate it to another person. Fehr, Herz and Wilkening (2013) similarly find that principals do not delegate decision rights to agents often enough in games where delegation results in higher monetary payoffs for both parties. Bartling, Fehr and Herz (2014) show that this underdelegation is driven by individuals assigning a positive value to decision rights *per se*. More specifically, they are willing to give up 16.7 percent of their expected payoff to retain control over their own payoff and the payoff of an agent they are matched with.

Our study is the first to separate power from autonomy and control and to show that preferences for power exist and are substantial, i.e., preferences for power are an important component of the intrinsic value of decision rights. Indeed, preferences for decision rights are a compound substance: when a principal retains decision rights she enjoys power, her ability to influence the outcomes of others, as well as control and autonomy,

[6] We are unable to conclusively determine the preference classes of about 11 percent of our subjects.

i.e., being able to influence her own payoff while being independent from the actions of others. In fact, in many experimental setups, including those in the above-mentioned papers, participants have to decide between having power, control, and autonomy, or having none of those.

In addition, our findings contribute to the corporate finance and delegation literatures that consider private benefits of decision-making as one of the main frictions in the principal-agent problem and in the optimal organizational design (e.g., Grossman and Hart (1986), Aghion and Bolton (1992), Hart and Moore (2005), Dessein and Holden (2017)). The theoretical literature has incorporated the non-pecuniary nature of private benefits. Hart and Moore (1995), for example, motivate their theory by claiming that "among other things, managers have goals, such as the pursuit of power" (p. 568). By their very nature, non-pecuniary private benefits are difficult to observe and even more difficult to quantify in a reliable way. Instead, the empirical literature has concentrated on measuring pecuniary private benefits by estimating the value of perquisites enjoyed by top executives (Demsetz and Lehn (1985), Dyck and Zingales (2004)). Dyck and Zingales (2004) find substantial evidence that good institutions and corporate governance can significantly curb the amount of monetary private benefits enjoyed by controlling shareholders. In the presence of preferences for power per se, however, even the best institutions would be unable to eliminate private benefit frictions.

Our results are also related to the literature on procedures versus outcomes. In strategic games, when evaluating decisions of others, individuals may base their assessments not only on outcomes but also on the procedures that lead to those outcomes. Indeed, past work has shown that including a third party in the decision-making process, changing the distance between a decision-maker and a recipient, varying the possibility of retribution and modifying the interpretation of motives and intentions leads individuals to evaluate outcomes differently. This is the case, for example, in Fershtman and Gneezy (2001), Coffman (2011), Bartling and Fischbacher (2011) and Orhun (2017). In our paper, we show that a large fraction of individuals care about procedures when it comes to how they *themselves* reach decisions concerning *others*, as opposed to how someone else acts towards them or others. This is the case even in the absence of strategic interactions, any possibility of retribution and in situations where beliefs regarding others' subsequent actions are irrelevant.

Finally, our findings have important methodological implications for inferring social preferences from individual choices. For example, Zizzo and Oswald (2001), Abbink and Sadrieh (2009), and Charness, Masclet and Villeval (2014) show that when people can choose by how much to decrease the payoffs of others, many of them are willing to sacrifice their own payoffs in order to "burn" other people's money. However, our study demonstrates that when individuals have preferences for power over the payoffs of others, a large fraction of them gladly exercise it, even though they do not attach any value to those payoffs per se. Our study reconciles results from these above papers with those studies that have shown that when people can only pick between two fixed options, where one of the options gives them less money but also destroys the payoffs of their partners, they behave in a much less malicious way (Charness and Rabin (2002), Chen and Li (2009)).

The remainder of the paper is organized as follows. In Section 2 we detail the Power Game and its experimental implementation. Section 3 outlines the theoretical framework behind our experimental design. Section 4 reports and discusses the experimental results. Section 5 concludes.

2. Experimental Design: The Power Game

2.1. **The Power Game.** We develop a new game, the "Power Game" and describe it here. The Power Game has two parts. At the beginning of Part I, players are randomly assigned a type, either A or B, with equal number of type As and type Bs. Types are fixed throughout the entire game. In the Power Game, only type A players make decisions.

Power Game, Part I: Part I comprises $N+1$ rounds. In each round, each type A player is randomly matched with a type B player. In round j, a price p_j is revealed to type A players who must then decide whether to pay it or not.

- If player A decides to pay p_j then the payoffs for both players are $(E_A - p_j, x^*_{Bj})$, where A chooses x^*_{Bj}, the payoff for player B, in the interval $[0, E_B]$.
- If player A decides not to pay p_j then the payoffs for both players are (x^*_{Aj}, E_A), where player A chooses x^*_{Aj}, her own payoff, in the interval $[0, E_A]$.[7]

Thus, a round consists of two stages: in the first stage, player A decides whether to pay p_j or not, and then, in the second stage, depending on her first stage decision, she chooses either her own or player B's payoff, i.e., either x_{Aj} or x_{Bj}. The values of E_A and E_B are known in advance and fixed throughout all the rounds. In the beginning of each round, for each player A, the price p_j is randomly and independently drawn from a discrete set $\mathcal{P}$ without replacement, and revealed to players before they make a decision on whether to pay it or not.

After all prices in the set $\mathcal{P}$ have been drawn, A players participate in a final round in Part I, round $N+1$. In this round, each type A player is given E_A as her payoff and asked to choose how much player B receives, still between 0 and E_B. In other words, here we

[7] See Section 2.2 for a detailed discussion on this design feature.

force all type A players to choose payoffs for type B players, as if they face a price of zero.

Power Game, Part II: Part II lasts for M rounds where $M \geq N$. In each round, player A decides between two payoff pairs: (x_A, x_B) and (x'_A, x'_B). N of the M rounds correspond to the first N rounds in Part I. These are player-specific as they depend on a player's decisions in Part I of the Power Game. More specifically, for each $p_j \in \mathcal{P}$:

- If in round j of Part I player A decided to pay p_j, then in the corresponding round of Part II, she decides between the following payoff pairs: (E_A, E_A) and $(E_A - p_j, x^*_{Bj})$, where x^*_{Bj} is the payoff she chose for player B in round j of Part I.
- If in round j of Part I player A decided not to pay p_j, then she chooses between $(E_A - p_j, E_A + 2p_j)$ and (x^*_{Aj}, E_A), where x^*_{Aj} is the payoff she chose for herself in round j of Part I. For rational subjects with standard preferences, this choice is effectively between $(E_A - p_j, E_A + 2p_j)$ and (E_A, E_A), since they always choose the maximum allowable on the interval $[0, E_A]$ in Part I.[8]

Whether or not a player paid p_j in round j of Part I, one of the payoff pairs she faces in the corresponding round of Part II is the pair she actually chose in Part I: $(E_A - p_j, x^*_{Bj})$ for players who paid and (x^*_{Aj}, E_A) for those who did not. The second option is an alternative payoff pair that the player could have chosen in round j of Part I but rejected: (E_A, E_A) if the player paid p_j and $(E_A - p_j, E_A + 2p_j)$ if she did not pay. (E_A, E_A) is an obvious choice in the former case since it is what a rational agent would have chosen if she had not paid. We adopt $(E_A - p_j, E_A + 2p_j)$ in the latter case because it is an efficient choice for all prices. Importantly, for each p_j a player encountered in Part I, in Part II she faces a choice between two payoff pairs, one of which is *identical in payoff distribution* to the pair that she actually selected in Part I, and the other is a pair she rejected.

In Part II, as in Part I, p_j is the cost to player A of moving away from (E_A, E_A). An important difference between paying in Part I and Part II is that in Part I a decision to pay leads to increased power over B's payoff. Indeed, if A pays in Part I, she extends her choice set for B's payoff from one element, E_A, to the whole interval $[0, E_B]$ and can select *any* number in that interval. The size of the choice set (in this case, interval $[0, E_B]$) should be of a little importance as long as by paying p, subjects can extend their choice set with respect to the other player's payoff. On the contrary, in Part II, paying does not lead to increased power since in either case, A's choice set for player B contains only one element.

The payoff pairs in the remaining $M - N$ rounds in Part II are chosen independently of Part I and correspond to other choices that may be of a separate interest to the researcher. See Section 2.2 for our choices.

[8]See Section 2.2 for further details.

2.2. **Experimental Implementation.** All our experimental sessions were conducted in March and April 2017 at the Laboratory for Economic Management and Auctions (LEMA) at the Pennsylvania State University using z-Tree software (Fischbacher (2007)). Subjects were recruited from the general undergraduate population and each subject participated in one session only. We conducted 16 sessions for a total of 292 subjects. Each session lasted at most 45 minutes and on average participants earned \$15. At the start of the experiment, subjects were randomly assigned a type: A or B. Subjects were told that throughout the entire experiment only Type A players' decisions would matter for payment and that types would remain fixed. Instructions for Part I were read out loud and afterwards all subjects participated in two practice rounds for Part I, where they could see what screens would look like if they did or did not decide to pay price p. In each round, each type A player was randomly matched with a type B player. Subjects moved from one round to the next when all subjects had completed the previous round. Instructions for Part II were handed out and read out loud after Part I was completed. Thus our subjects were not aware of the contents of Part II when they were making their Part I decisions. After the end of Part II, subjects filled out a questionnaire where we asked them what motivated their choices, as well as demographic and education information. Full instructions are available in Appendix A. The questionnaire is available in Appendix B.

<u>Type assignment:</u> Types were assigned at the beginning of the experiment. However, the subjects were not told what type they were, but were told to make decisions as if they were type A players. If their *true* type was B, none of their decisions would matter for payment. If their true type was A, then one of their decisions, randomly selected, would matter. Thus, regardless of one's true type, it was in one's best interest to make decisions as if one were a type A player.

<u>Parameter values in Part I:</u> We used the following parameter values in Part I. The price set $\mathcal{P}$ contained 9 distinct prices ranging from \$0 to \$2, in increments of 25 cents: $\mathcal{P} = \{\$0, \$0.25, \$0.50, ... \ \$1.50, \$1.75, \$2\}$. Thus, subjects played 9 rounds where prices were drawn without replacement from $\mathcal{P}$. Each price was randomly and independently drawn for each subject in each round. Subjects then played round 10 in which they were forced to choose the payoff for player B, as if price p_{10} was equal to 0. We used $E_A = \$12.30$ and $E_B = \$16.30$.

Subjects were not aware of the contents of $\mathcal{P}$, they were simply told the price would vary from round to round. If A decided to pay, she would receive $\$12.30 - p$ as her payoff and she would obtain the right to choose the payoff for B, and could choose any number between \$0 and \$16.30 (in increments of 5 cents).[9] If A decided not to pay, B would receive \$12.30 and A could choose how much to give to herself, between \$0 and \$12.30 (in increments of 5 cents).

Before starting Part I of the Power Game, subjects participated in two practice rounds. In those rounds, they were shown the screens that paying and not paying would lead to. Thus, they could familiarize themselves with the game and satisfy any curiosity regarding what paying or not would lead to in terms of screen display.

A few elements of our design are worth elaborating upon.

- Subjects choose their own payoff when they do not pay p. This is done for several reasons. First, in this way, in all rounds of Part I, all subjects make similar decisions: after deciding to pay p or not, they have to select how much to give to themselves or to the subjects they are matched with. Thus, this design element keeps anonymity fuller: no subject can infer whether another has decided to pay or not since all subjects type after making their first stage decisions. Second, we mitigate any experimenter demand effect where subjects might decide to pay in Part I simply because it is the only option with a subsequent action. Finally, it minimizes decisions to pay that would simply be due to boredom.[10]
- Making $E_A < E_B$. The advantage is three-fold. First, this allows us to explore a broad range of social preferences. Second, subjects who have preferences for power are not limited in how they can exercise it: they are not constrained to increase or decrease B's payoff relative to what B would receive if A does not pay. Finally, we can better study the interaction of power and social preferences.
- Subjects are not told what type they are. This design feature allows us to collect decisions from all our subjects since they all behave as if they were type A players, as opposed to revealing types and only collect data from half of the subjects in each session. Our instructions carefully describe this design element and subjects are emphatically told that they should act as type A players, since if their true type were B none of their decisions would matter (see Appendix A for the instructions). In one of the rounds of Part II we directly test whether subjects understood their roles and find strong evidence that they did. In that specific round, all subjects are faced with a choice between two payoff pairs, $(12.30, 9.60)$

[9] Strictly speaking, by paying p, player A acquires a right to choose any payoff for B among 327 options. She increases the size of her choice set for B's payoff from 1 element (\$12.30 only) to 327 elements (\$0, \$0.05, \$0.10, ..., \$16.25, \$16.30).

[10] There is however one potential drawback to this design choice. Subjects may enjoy choosing for themselves, i.e., they may enjoy control. This would mean that subjects enjoy choosing \$12.30 for themselves more than receiving \$12.30 when that amount is pre-specified. If this is the case, we might underestimate the fraction of people who enjoy power if they have stronger preferences for control than for power. Given the advantages of letting type A players choose their own payoffs, we proceeded with that design feature.

and $(9.60, 12.30)$. Which pair appears on the left or on the right of the screen is randomly decided (as for all rounds in Part II), yet 96% of our subjects choose $(12.30, 9.60)$. Had there been any doubt on who to make decisions for, the fraction choosing the latter would have been higher.

- Subjects are told that throughout the entire experiment, types are fixed and only A players make decisions that matter for payment. These design elements minimize the possibility that subjects' decisions in Part I are motivated by their belief that those decisions may be rewarded or used against them in some way by other subjects in Part II.
- Unordered prices: p is randomly drawn from $\mathcal{P}$. In some experimental designs, the experimenter restricts the choices of subjects so that they appear rational and "well behaved", e.g., such that all subjects have cutoff strategies. In our context this would mean imposing that as soon as for some price a subject decides not to pay, we force that the rest of her decisions be "not pay" for any price greater than that first price. Another way to "encourage" well-behaved choices is to offer an ordered list of prices to the subjects. We however let price p be randomly drawn from $\mathcal{P}$ and ask subjects to make decisions for all prices in $\mathcal{P}$, regardless of past behavior. We do so for two reasons. First, we are able to identify the subset of subjects who *are* well behaved and conduct several analyses: using those subjects only and using the entire sample. We can evaluate whether our results depend on the kind of subjects we are considering.[11] Second, random price order ensures that our results are not driven by any order effects.[12]

Parameter values in Part II: Part II consisted of 20 rounds where subjects decided between two payoff pairs. Which payoff pair was presented on the left or on the right of the screen was randomly determined for each subject in each round. 9 rounds were subject-specific and 11 rounds were identical for each subject. The order of rounds was random for each subject.

In Part II, the 9 subject-specific rounds depended on a particular subject's decisions over the first 9 rounds of Part I. Specifically, subjects decided between the payoff pair they chose in Part I and a pair that was available but rejected:

[11]In the main text we focus on well-behaved subjects, and in Appendix C we re-do the analysis including subjects who skipped one price and in Appendix D include all the subjects. We show that our results are unchanged across the different samples.

[12]Relatedly, Brown and Healy (2016) show that designs in which choices are all listed together on one screen are not incentive compatible, whereas designs in which choices are randomly presented are.

Table 1. Decision Problems in 11 Rounds of Part II.

Decision[a]	First Option[b]	Second Option
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$
PT1	$(10.10, 5.20)$	$(9.10, 9.10)$
PT2	$(12.30, 5.10)$	$(10.10, 12.30)$
PT3	$(12.30, 9.60)$	$(9.60, 12.30)$
PT4	$(12.30, 7.80)$	$(7.80, 5.40)$
PT5	$(6.15, 6.15)$	$(0.00, 0.00)$

[a]These rounds were presented among 20 rounds of Part II in random order for each subject.
[b]What option was presented on the left or on the right of the screen was randomly determined independently for each decision problem and for each subject.

- If a subject decided to pay p_j and chose x_{Bj} in round j of Part I, she had to choose between the following payoff pairs in the corresponding round in Part II: $(12.30 - p_j, x_{Bj})$ and $(12.30, 12.30)$.
- If a subject decided not to pay p_j and chose x_{Aj} in round j of Part I, she had to choose between the following payoff pairs in the corresponding round in Part II: $(x_{Aj}, 12.30)$ and $(12.30 - p_j, 12.30 + 2p_j)$.

The remaining 11 rounds were identical for each subject. In six of those rounds, the values for the payoffs pairs were inspired by Charness and Rabin (2002)[13] and re-scaled such that the order of magnitudes for payoffs was similar to the values stemming from Part I; see decisions CR1-CR6 of Table 1. Other decision problems were chosen to be similar to some of the problems in Charness and Rabin (2002) but to allow for different trade-offs between the payoffs of players A and B; see decisions PT1 and PT2 in Table 1. Decision problem PT3 was designed to check whether subjects understood that they were to act as type A players. Finally, problems PT4 and PT5 were chosen to serve as "sanity checks" in our analysis (for more details see Section 4.3.2).

[13]See two-person dictator games, Table 1, p. 829.

3. Theoretical Framework

In this section, we derive a set of theoretical predictions for Parts I and II of the Power Game for individuals with different preference classes. We think of individual preferences varying along two dimensions. The first is whether an individual non-trivially incorporates other players' payoffs in her utility function. The second is whether she derives utility from having power over payoffs of others.[14] Thus, we consider the following four types of preferences: standard selfish preferences, social preferences, power preferences, and, since power and social preferences are not mutually exclusive, preferences that have both social and power components.

We start with specifying player A's utility function in a general form: $U_A = U(x_A, x_B, \lambda)$, where x_A is A's own payoff, x_B is B's payoff, and λ is a parameter indicating the amount of power that A has over B's payoff. Without loss of generality we normalize λ to zero when B's potential payoff is pre-specified, i.e., when A receives no additional power from her choice, as is the case in all Part II rounds as well as if she doesn't pay in Part I. Similarly, we impose $\lambda = 1$ when A chooses B's payoff from the interval $[0, E_B]$. We make the following assumptions about A's utility function.

Assumption 1. $U(x_A, x_B, \lambda)$ *is continuous in all three arguments.*

Assumption 2. *For all x_B and all λ, $U(x_A, x_B, \lambda)$ is strictly increasing in x_A.*

Assumption 3. $U(x_A, x_B, \lambda) = V(x_A, x_B) + f(\lambda)$, *where $f(\lambda) = 0$ for players who do not care about power and $f(\lambda)$ is strictly increasing in λ for power-hungry players.*

Assumption 1 simply states that player A's utility function is continuous for all values of its arguments x_A, x_B, and λ. In Assumption 2 we impose that all else equal, A's utility function is strictly increasing in her own payoff, but we allow for a large set of preferences as they relate to other players' payoffs.[15] Assumption 3 states that preferences for power enter the utility function in an additively separable manner and they are monotonic.

Before we derive predictions about willingness to pay in Parts I and II of the Power Game for different preference classes, we show that all players' demand functions are well-behaved.

Lemma 1. *Player A has a well-behaved demand function in both Parts of the Power Game.*

[14] In the main text, we focus on players who derive positive or no utility from power as opposed to deriving negative utility from it. We nevertheless, when appropriate, describe how to modify the theory for such subjects.

[15] Examples of utility functions that incorporate social preferences can be found in Rabin (1993), Levine (1998), Fehr and Schmidt (1999), Bolton and Ockenfels (2000), Charness and Rabin (2002), Falk and Fischbacher (2006), Cox, Friedman and Gjerstad (2007), Cox, Friedman and Sadiraj (2008), and Chen and Li (2009).

Proof. Here we concentrate on behavior in Part I of the Power Game as the proof for Part II is similar and therefore omitted here for the sake of brevity. We show the existence of $\bar{p}_I$ in two steps. First, we prove that if a subject pays p', then she pays for all $p < p'$. We then prove that if a subject does not pay p'', then she does not pay for any $p > p''$. These steps suffice to show that demand functions are well-behaved.

Suppose A pays p'. Since A pays p', we must have that $\mathrm{U}(E_A - p', x_B^*, 1) \geq \mathrm{U}(E_A, E_A, 0)$, where x_B^* is what A chooses for B. Since $\mathrm{U}(x_A, x_B, \lambda)$ is strictly increasing in x_A, for any $p < p'$, we have that $\mathrm{U}(E_A - p, x_B^*, 1) > \mathrm{U}(E_A - p', x_B^*, 1) \geq \mathrm{U}(E_A, E_A, 0)$. In other words, at price p player A is better off choosing x_B^* than not paying. Thus, for all $p < p'$, A pays and A's demand function is well-behaved for all $p < p'$.

Suppose A does not pay p''. Then, $\mathrm{U}(E_A - p'', x_B^{**}, 1) \leq \mathrm{U}(E_A, E_A, 0)$, where x_B^{**} is what A chooses for B at $p = p''$. By monotonicity of $\mathrm{U}(x_A, x_B, \lambda)$ in x_A, for any $p > p''$ we have, for all x_B, that $\mathrm{U}(E_A - p, x_B, 1) < \mathrm{U}(E_A - p'', x_B^{**}, 1)$. Thus, for all $p > p''$, player A does not pay and A's demand function is well-behaved for all $p > p''$. □

Lemma 1 shows in each Part of the Power Game players follow one of three paying behaviors: (1) a player never pays a positive price; (2) a player pays up until a cutoff price but does not pay for any price above it; or (3) a player pays at all prices. Note that to guarantee a finite cutoff price, additional constraints are needed: $f(\lambda)$ is bounded from above or A faces a budget constraint are each individually sufficient. Heretofore, we call player A's willingness to pay in Part I $\bar{p}_{\mathrm{I}}$ and call A's willingness to pay in Part II $\bar{p}_{\mathrm{II}}$.

We now turn to deriving the behavior of subjects with different classes of preferences. We begin with a player who has "standard" (completely selfish) preferences and who does not derive any utility from having power: $\mathrm{U}(x_A, x_B, \lambda) > \mathrm{U}(x'_A, x'_B, \lambda')$ for all $x_A > x'_A$, all x_B and x'_B, and all λ and λ'. Proposition 1 states that a player with standard preferences always chooses to maximize her own payoff. Such a player never pays positive prices in either part of the Power Game.

Proposition 1. *For a player with standard preferences, $\bar{p}_I = \bar{p}_{II} = 0$.*

Proof. In Part I, if player A pays $p > 0$, then her payoff is equal to $E_A - p$. If she does not pay, her payoff is E_A. Since she only cares about her own payoff, she does not pay for any $p > 0$ and pays for $p < 0$. Thus, $\bar{p}_{\mathrm{I}} = 0$, and similarly $\bar{p}_{\mathrm{II}} = 0$. □

Next, we consider the behavior of players with non-standard preferences, i.e., players with social preferences or power preferences or both. We make two additional assumptions about the utility function of such players. We assume that if a player is indifferent between two options, she always chooses the one that gives her the highest monetary payoff. In particular, if a player pays p, this represents a strict preference ordering. We include this assumption for convenience and it is innocuous vis-à-vis our results. Finally, we assume that if A has social preferences, then her utility is highest at a single point in the interval $[0, E_B]$. That is, for a given x_A there are no two different payoffs for B that give A that highest utility. Assumptions 4 and 5 formalize these notions.

Assumption 4. *If $x_A > x'_A$ and $\mathrm{U}(x_A, x_B, \lambda) = \mathrm{U}(x'_A, x'_B, \lambda')$, then a player chooses the (x_A, x_B, λ) option.*

Assumption 5. *For a player with social preferences, for any x_A and λ, there exists a unique x^*_B that maximizes her utility, i.e.,* $\underset{x_B \in [0,E_B]}{\operatorname{argmax}}\ \mathrm{U}(x_A, x_B, \lambda) = \{x^*_B\}$.

Let us first consider players who have social preferences and no power preferences. Such a player incorporates B's payoff in her utility in a non-trivial manner, but because she is indifferent towards power, her utility from a particular allocation (x_A, x_B) is not affected by how that allocation is obtained. In other words, this player's utility is the same whether she chooses a particular payoff for player B from the interval $[0, E_B]$ or whether it is exogenously given. For this player $\mathrm{U}(x_A, x_B, \lambda) = \mathrm{U}(x_A, x_B, \lambda')$, for any λ λ', that is, for all λ, $f(\lambda) = 0$.

Social preferences that satisfy our assumptions can be divided into two categories that are mutually exclusive and together comprise the entire set of social preferences:

(1) Social preferences where player A maximizes her utility by choosing something other than E_A for player B when $p = 0$. That is, there exists $x^{\circledast}_B \neq E_A$ such that A strictly prefers $(E_A, x^{\circledast}_B, \lambda)$ to (E_A, E_A, λ), i.e., $E_A \neq \underset{x_B \in [0,E_B]}{\operatorname{argmax}}\ \mathrm{U}(E_A, x_B, \lambda)$.[16]

(2) Social preferences where player A maximizes her utility by choosing E_A for player B when $p = 0$. That is, player A strictly prefers (E_A, E_A, λ) to (E_A, x_B, λ) for any $x_B \neq E_A$, i.e., $E_A = \underset{x_B \in [0,E_B]}{\operatorname{argmax}}\ \mathrm{U}(E_A, x_B, \lambda)$.[17]

In Proposition 2 and Corollary 1 we derive predictions regarding the paying behavior of players with social preferences and no power preferences in the Power Game.

Proposition 2. *For a player with social preferences and no power preferences, $\bar{p}_I > 0$ if and only if she does not choose E_A for player B when $p = 0$. In Part II, player A has the same paying behavior as in Part I: $\bar{p}_I = \bar{p}_{II}$.*

Proof. Let us start by showing that if a player does not choose E_A for B when $p = 0$, then there exits some positive price $p' > 0$ such that A pays p'.

[16]Competitive or spiteful preferences and social-welfare preferences as in Charness and Rabin (2002) and Cox, Friedman and Gjerstad (2007) are in this category.

[17]Preferences for equality in Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) are example that fall into this category of preferences.

Let $x_B^{\circledast}$ be what A chooses for B when $p = 0$, i.e., $x_B^{\circledast} = \underset{x_B \in [0,E_B]}{\operatorname{argmax}}\ \mathrm{U}(E_A, x_B, 1)$, which is a singleton by Assumption 5. Since A does not choose E_A for B at $p = 0$, we know that $x_B^{\circledast} \neq E_A$, and thus $\mathrm{U}(E_A, x_B^{\circledast}, 1) > \mathrm{U}(E_A, E_A, 0)$. By continuity and monotonicity of $\mathrm{U}(x_A, x_B, \lambda)$ in x_A, there exists $p' > 0$ such that $\mathrm{U}(E_A, x_B^{\circledast}, 1) > \mathrm{U}(E_A - p', x_B^{\circledast}, 1) > \mathrm{U}(E_A, E_A, 0)$. Thus, there exists $p' > 0$ such that A pays p'. By Lemma 1, since A pays a positive price at least once, her willingness to pay $\bar{p}_{\mathrm{I}}$ is strictly positive.

Now we show the opposite direction: if $\bar{p}_{\mathrm{I}} > 0$, then A does not choose E_A for B at $p = 0$. Since $\bar{p}_{\mathrm{I}} > 0$, by continuity and monotonicity of $\mathrm{U}(x_A, x_B, \lambda)$ in x_A, there exists some positive price $p < \bar{p}_{\mathrm{I}}$, such that $\mathrm{U}(E_A - p, x_B^*, 1) > \mathrm{U}(E_A, E_A, 0)$, where $x_B^* = \underset{x_B \in [0,E_B]}{\operatorname{argmax}}\ \mathrm{U}(E_A - p, x_B, 1)$ and $x_B^* \neq E_A$. By monotonicity, we have that $\mathrm{U}(E_A, x_B^*, 1) > \mathrm{U}(E_A - p, x_B^*, 1) > \mathrm{U}(E_A, E_A, 0)$. In other words, at $p = 0$, player A is better off when she chooses x_B^* as player B's payoff than when she chooses E_A. Thus, at a price of 0, player A pays and does not choose E_A for player B.

The second statement of the proposition follows directly from the fact that $\mathrm{U}(x_A, x_B, \lambda) = V(x_A, x_B) + f(\lambda)$ and that A doesn't care about power, i.e., $f(\lambda) = 0$ and for all x_A, x_B, $\mathrm{U}(x_A, x_B, 1) = \mathrm{U}(x_A, x_B, 0)$. Thus, the Part I logic holds for Part II and $\bar{p}_{\mathrm{II}} = \bar{p}_{\mathrm{I}}$. $\square$

Proposition 2 shows that players whose preferences fit within a large class of social preferences, e.g., social-welfare preferences or competitive preferences, have a strictly positive willingness to pay in Part I of the Power Game. Their willingness to pay, $\bar{p}_{\mathrm{I}}$, depends on the strength of their preferences with respect to the payoff of player B as well as on the available payoff options for B, i.e., the interval $[0, E_B]$. Moreover, such players are willing to pay the same positive price in Part II, i.e., $\bar{p}_{\mathrm{I}} = \bar{p}_{\mathrm{II}}$.

Corollary 1. *For a player with social preferences and no power preferences, $\bar{p}_I = 0$ if and only if such a player chooses E_A for player B when $p = 0$. Moreover, $\bar{p}_{II} = \bar{p}_I = 0$.*

Corollary 1 is the transpose of Proposition 2 and so is directly implied by it. One direct implication of Corollary 1 is that players with preferences for equality should never pay in either Part of the Power Game.

We now consider players who like power. If a player enjoys being able to choose payoffs for others, her utility has to incorporate not only final payoffs, but also whether those payoffs are attained via increased power: for $\lambda > \lambda'$, $\mathrm{U}(x_A, x_B, \lambda) > \mathrm{U}(x_A, x_B, \lambda')$ for all x_A and x_B. In other words, for such players, $f(\lambda)$ is strictly increasing in λ. Such players may or may not have social preferences. We start with those who do not.

Proposition 3. *For a player with power preferences and no social preferences, $\bar{p}_I > 0$ and $\bar{p}_{II} = 0$.*

Proof. Since player A derives a positive utility from having power and does not have social preferences, then in Part I of the Power Game, for all x_B, $U(E_A, x_B, 1) > U(E_A, E_A, 0)$. By continuity and monotonicity of $U(x_A, x_B, \lambda)$ in x_A, there exists $p' > 0$ such that $U(E_A, x_B, 1) > U(E_A - p', x_B, 1) > U(E_A, E_A, 0)$, and A pays $p' > 0$. By Lemma 1, we conclude that $\bar{p}_{\mathrm{I}} > 0$.

In Part II of the Power Game, paying does not lead to any increase in power for A, since in either case the potential payoff for B is fixed. Therefore, for any $p > 0$ and any x_B, x'_B, $U(E_A, x_B, 0) > U(E_A - p, x'_B, 0)$. Thus, A never pays a positive price in Part II: $\bar{p}_{\mathrm{II}} = 0$. □

Proposition 3 states that a player with power preferences and no social preferences is willing to pay positive prices in Part I of the Power Game. In Part II however she instead chooses the payoff-maximizing option and never pays a positive price.

Finally, we consider subjects who have both social and power preferences.

Proposition 4. *For a player with power and social preferences, $\bar{p}_I > 0$ and $0 \leq \bar{p}_{II} < \bar{p}_I$.*

Proof. Let's first consider the social component of A's utility function. By Assumption 3, for a given payoff of player A, there is an optimal allocation for B that is independent of her preferences for power. That is, for every λ, λ', $x_B^* = \underset{x_B \in [0, E_B]}{\operatorname{argmax}} U(x_A, x_B, \lambda) = \underset{x_B \in [0, E_B]}{\operatorname{argmax}} U(x_A, x_B, \lambda') = \underset{x_B \in [0, E_B]}{\operatorname{argmax}} V(x_A, x_B)$.

In particular, if her payoff is $E_A - p$, the allocation for B that maximizes her utility is the same in Part I and in Part II. By Proposition 2 and Corollary 1, a player with a social component to her utility is willing to trade-off $p_s \geq 0$ to maximize her utility and obtain the $(E_A - p_s, x_B^*)$ allocation instead of the (E_A, E_A) allocation, i.e., $V(E_A - p_s, x_B^*) \geq V(E_A, E_A)$, where $x_B^* = \underset{x_B \in [0, E_B]}{\operatorname{argmax}} V(E_A - p_s, x_B)$.

In a setting of increased power, her choice for B is identical and her utility is then $V(E_A - p_s, x_B^*) + f(1)$. Because she enjoys power, that utility is strictly greater than $V(E_A - p_s, x_B^*) + f(0)$. By continuity and monotonicity of $V(x_A, x_B)$ in x_A, there exists $p' > p_s$ such that $V(E_A - p_s, x_B^*) + f(1) > V(E_A - p', x_B^*) + f(1) > V(E_A - p_s, x_B^*) \geq V(E_A, E_A)$. Equivalently, $U(E_A - p_s, x_B^*, 1) > U(E_A - p', x_B^*, 1) > U(E_A - p_s, x_B^*, 0) \geq U(E_A, E_A, 0)$. Thus, A is willing to pay more to implement both her social and power preferences than to implement her social preferences only. As a result A is willing to pay more in Part I than in Part II. □

Proposition 4 states that if a player enjoys power and has social preferences, she is willing to sacrifice a larger fraction of her payoff in order to both obtain power and implement her social preferences as in Part I than in order to only implement her social preferences, as in Part II.

3.1. **Experimental Predictions and Empirical Identification of Preference Classes.** In our theory, we have established the correspondence between paying behavior in the Power Game and preference classes. Before we outline how our theoretical predictions translate to the empirical identification of preference classes, we discuss two caveats.

The first caveat regards the experimental implementation of the Power Game. Since in the experiment subjects face a menu of prices in increments of 25 cents, it is possible that we are unable to observe the willingness to pay for those subjects with relatively weak power and/or social preferences. Indeed, their willingness to pay may be below 25 cents. In this case, we may underestimate the fraction of people with power and/or social preferences as these may instead appear to us as having standard preferences.

Next, while players with power preferences are clearly identified in the Power Game, for a set of social preferences, isolating players who exclusively care about power from those who care about power *and* have those specific social preferences poses a challenge. Players with those social preferences choose \$12.30 as B's payoff when their own payoff is \$12.30, i.e., when $p = 0$ and in Round 10 of Part I. Such players pay positively in Part I but not in Part II, exactly like players who have power preferences only.[18] We choose to categorize players who only pay positively in Part I as having power preferences only, with the caveat that a small fraction may in fact also have social preferences.[19] Importantly, this choice does not impact the overall fraction of subjects who have a power component to their preferences.

TABLE 2. Empirical Identification of Preference Classes.

Preference class	$\bar{p}_{\mathrm{I}}$	$\bar{p}_{\mathrm{II}}$
Standard[a]	0	0
Power +	$\bar{p}_{\mathrm{I}} > 0$	0
Social Preferences	$\bar{p}_{\mathrm{I}} > 0$	$\bar{p}_{\mathrm{II}} = \bar{p}_{\mathrm{I}}$
Social Preferences & Power +	$\bar{p}_{\mathrm{I}} > 0$	$\bar{p}_{\mathrm{II}} < \bar{p}_{\mathrm{I}}$
Social Preferences & Power −	$\bar{p}_{\mathrm{I}} \geq 0$	$\bar{p}_{\mathrm{II}} > \bar{p}_{\mathrm{I}}$
Unclassified	.	Any

[a]Players with $\bar{p}_{\mathrm{I}} = \bar{p}_{\mathrm{II}} = 0$ might have social preferences such that they choose \$12.30 as B's payoff when their own payoff is \$12.30, i.e., when $p = 0$ and in Round 10 of Part I. However, in our empirical analysis, we find that none of our subjects do so.

Table 2 summarizes the correspondence between preference class and paying behavior across Parts I and II of the Power Game in our experimental setup. We expect subjects

[18]See Corollary 1 and Proposition 4.

[19]Empirically, the fraction of subjects who may have power preferences and those social preferences is at most 3% of our sample, since this is the proportion of subjects who choose \$12.30 as B's payoff when their own payoff is \$12.30. This small percentage represents an upper-bound on the fraction of subjects who may have this type of social preferences, insomuch as \$12.30 is a rather obvious focal point.

with standard preferences to never pay positive prices in either Part I or Part II. Subjects with social and/or power preferences pay strictly positive prices in order to choose the payoffs for others. Note that in both cases subjects' willingness to pay depends on the strength of their power and social preferences relative to their own payoff. While subjects with power or social preferences behave similarly in Part I, they are different in Part II: subjects with social preferences pay the same price in Parts I and II, whereas subjects with power preferences never pay positive prices in Part II. Subjects who have social preferences and value power are willing to pay higher prices in Part I than in Part II, but will still pay positively in Part II. If subjects have social preferences but dislike power, they pay more in Part II than in Part I. Finally, subjects who never pay in Part I, including at a price of zero, may be motivated by different factors that we are unable to separate with our experimental design. These subjects may have standard preferences and make a random decision on whether to pay at a price of zero or not. They may instead have preferences for equality when A herself receives \$12.30 and since both paying and not paying at zero leads to both players receiving \$12.30 they may choose not to pay. Finally, these subjects may dislike power and refuse to pay for it, even at a price of zero.

4. Experimental Results

In our analyses, we restrict our sample to those subjects who prefer to earn more rather than less, all else equal. That is, we use Assumption 2 in Section 3 as a guiding principle. Subjects violate Assumption 2 if in Part I they decide not to pay and choose a payoff for themselves that is lower than \$12.30. Out of our initial 292 subjects, 242 always give themselves \$12.30 and 16 subjects give themselves \$12 or more. We have 6 subjects who give themselves less than \$12 at least once, 13 do that twice, and 15 do that three or more times. For our analyses, we consider those 258 subjects, or 88.4% of our original sample, who always give themselves at least \$12.[20,21]

4.1. **Preferences for Power: the Aggregate Level.**

4.1.1. *Demand for choosing payoffs of others in Part I.* We begin by exploring the relationship between prices and subjects' decisions to pay for the right to choose precise payoffs of others in Part I. Figure 1 presents the fraction of subjects who agree to pay in Part I for

[20]We retain the 16 subjects that give themselves between \$12 and \$12.30, under the assumption that some subjects may have a preference for not receiving multiple coins as part of their final cash payment. For example, out of those 16 subjects, 11 subjects consistently give themselves exactly 12 dollars and three subjects choose 12 dollars and a quarter as their payoff.

[21]All our conclusions are robust to including all subjects as well as only including subjects who strictly adhere to Assumption 2.

each given price, including the price of zero.[22] Two features deserve emphasis. First, the fraction of subjects who pay to choose others' payoffs is decreasing in price, or in other words, the demand function is downward sloping. Second, at the price of $0, over 90% of our subjects decide to choose the payoff for player B. Both elements show that our subjects understand that there is a real cost to choosing the other player's payoff, which is perhaps not surprising since they were explicitly asked whether they were willing to "pay p" in order to choose the payoff for B. The existence of an aggregate downward-sloping demand function shows that the preferences for choosing payoffs of others are consistent and well-behaved at the aggregate level.

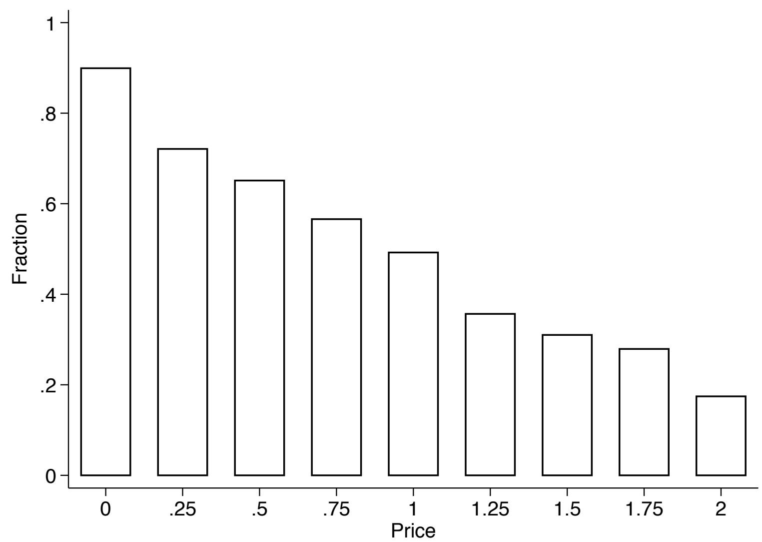

FIGURE 1. Fraction of subjects who pay to choose payoffs of others in Part I by price, in USD.

In Table 3 we present additional evidence on the participants' decisions to pay depending on price. For each subject, we calculate how many times he or she decides to pay for each of the nine prices from the set $\mathcal{P}$, which ranges from $0 to $2, in increments of 25 cents. The overwhelming majority of subjects, just over 80%, decide to pay a positive price at least once. Recall that before the start of Part I, subjects go through two practice rounds in which they are able to satisfy any curiosity regarding what such a choice would lead to in terms of screen display. Further, 70.2% of our participants pay at least twice and about 10.5% pay at all prices, including the maximum offered price of $2, which represents a loss of over 16% of what they could have earned had they not paid. In fact, the median number of strictly positive payments is 3, and the mean is slightly higher at 3.6.

[22]Recall that $\mathcal{P}$ includes the price of zero, and when faced with this price subjects still have to decide whether to pay it or not.

TABLE 3. Fraction of subjects paying to choose payoffs of others in Part I.

	Subjects	Percentage
Never paid, including at price of zero	15	5.8%
Paid only at price of zero	35	13.6%
Paid a positive price at least once	208	80.6%
Paid a positive price at least twice	181	70.2%
Paid at all prices, including at price of zero	27	10.5%
All subjects	258	100.0%

4.1.2. *Allocations chosen in Part I.* In Part I of the Power Game, our 258 subjects make 2,322 different decisions over the courses of the 9 prices they face. Note that if a subject decides not to pay, she effectively chooses the $(12.30, 12.30)$ allocation. This happens in 1,174 (or 50.56%) of those decisions. Here we concentrate on the remaining 1,148 (or 49.44%) of the allocations, when subjects choose to pay p, i.e. $(12.30 - p, x_B^*)$. Thus, we can depict those chosen allocations on the $x_A x_B$ plane and do so in Figure 2.

Figure 2 clearly demonstrates that there is substantial heterogeneity in terms of the payoff allocations chosen by our subjects. The surface of each circle is proportional to the number of subjects who choose a specific allocation. For example, when price p is 25 cents, 97 of our subjects decide to pay and to give \$16.30 to B, effectively choosing the $(12.05, 16.30)$ allocation, while when the price is 2 dollars, only 18 subjects choose the $(10.30, 16.30)$ allocation. All the allocations lying above the downward-sloping solid line are efficient in that player A pays less money than she gives to B beyond \$12.30. More formally: $x_B > 12.30 + p$. All the allocations below the upward-sloping dashed line are competitive[23] in that player A decreases the payoff of B by more than what she pays. More formally: $x_B < 12.30 - p$. In other words, in these cases player A is willing to slightly decrease her own payoff to decrease the payoff of B even further. In terms of allocation distribution, 66.1% are efficient, and the most efficient allocation of $(12.30 - p, 16.30)$ comprises 51.9% of the ones our subjects are willing to pay for. Surprisingly, 25.2% of the allocations are competitive, i.e., the subjects give less to B than they receive themselves. The remaining 8.7% of the allocations cannot be attributed to either category.[24]

[23]We adopt this terminology from Charness and Rabin (2002).

[24]Note that we use a strict relationship to attribute allocations to being efficient or competitive ($x_B > 12.30 + p$; $x_B < 12.30 - p$). If we allow for equality, then the percentages of efficient, competitive, and unclassified allocations are 67.5%, 26.1%, and 7.3% respectively.

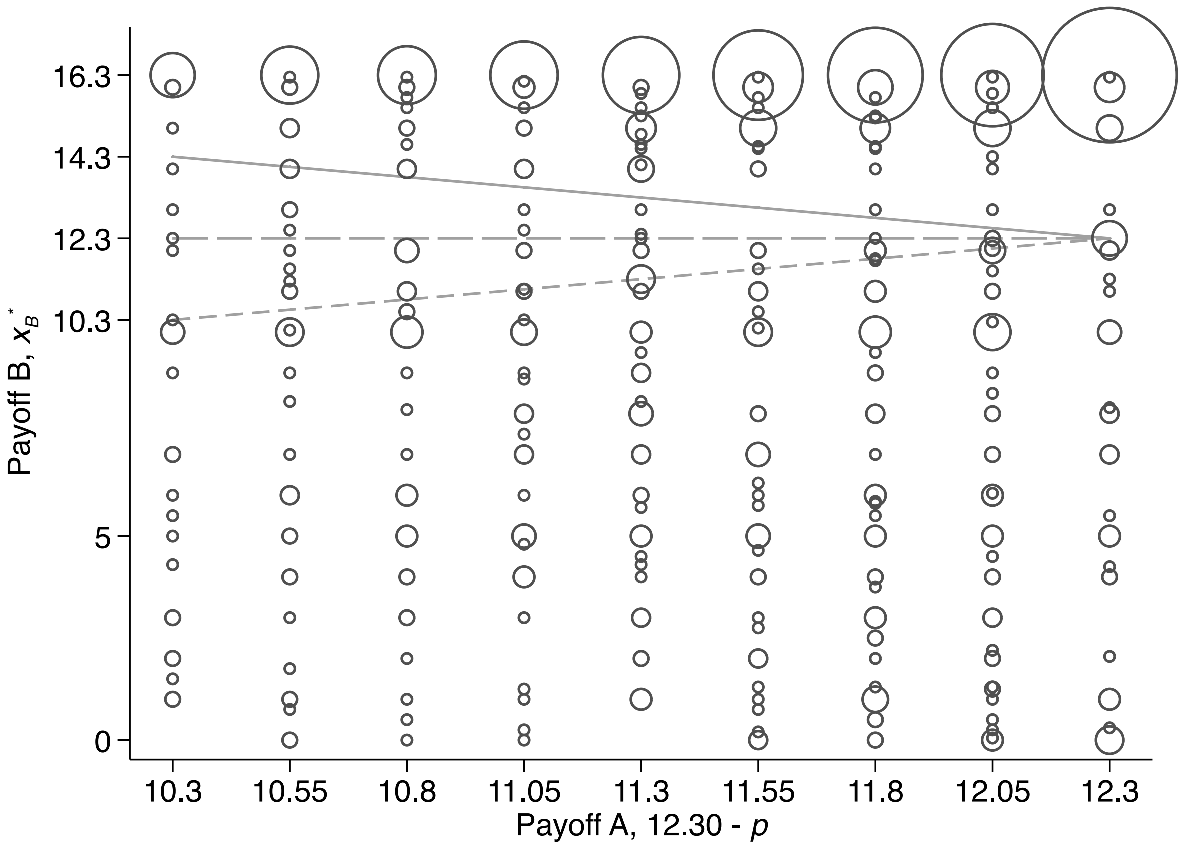

FIGURE 2. Payoff allocations $(12.30 - p, x_B^*)$ chosen by subjects in Part I.

In Table 4 we aggregate observations at the subject level. For each subject, we compute the average and maximum prices paid as well as the average amount given to player B across all decisions. In Panel A, we present the above statistics for all decisions, including at a price of zero, while in Panel B we do so, but only for decisions done at positive prices. Among our 258 subjects, 243 pay at least once, including at zero, and for them the median maximum price paid is \$1.25. In other words, the majority of those who pay are willing to give up substantial amounts to obtain the right choose the precise payoff for B players. Obviously, conditional on paying positive prices, the median maximum price paid is even higher. It reaches \$1.50 and constitutes 16% of the subjects' potential payoffs. Thus, not only are subjects willing to pay in order to choose B players' payoffs (as seen in Table 3), but they are also willing to pay relatively large amounts in order to do so.

Additionally, there is substantial variation among subjects in terms of the payoff chosen for player B. In both Panels A and B, the minimum average amount is zero and the maximum is \$16.30. That is, at least one subject always gives nothing, and one subject always gives the maximum allowable amount to player B. In addition, the distributions

TABLE 4. Subjects' behavior in Part I.

	Subjects	Mean	St.Dev.	Min.	p25	Median	p75	Max.
Panel A: At all prices								
Mean price paid	243	0.54	0.34	0.00	0.25	0.62	0.88	1.21
Max. price paid	243	1.11	0.71	0.00	0.50	1.25	1.75	2.00
Mean payoff for B	243	13.46	4.37	0.00	11.22	16.20	16.30	16.30
Panel B: Only at positive prices								
Mean price paid	208	0.79	0.31	0.25	0.50	0.85	1.00	2.00
Max. price paid	208	1.30	0.58	0.25	0.75	1.50	1.75	2.00
Mean payoff for B	208	12.85	4.70	0.00	10.10	15.65	16.30	16.30

shows that most participants tend to be generous rather than petty, and they tend to be especially nice when it costs them nothing. In fact, 117 out of 243 and 86 out of 208 subjects always give \$16.30 to player B, for all prices and strictly positive prices, respectively.

In Part I, A may decide to pay because she has social preferences or because she enjoys power. Thus, Part I of the Power Game cannot, on its own, clearly and cleanly distinguish between social preferences and preferences for power. Below, in Section 4.1.3 we address this issue and show that in fact, preferences for power *per se* are prevalent.

4.1.3. *The Power Game: Part I versus Part II behavior.* In this section we analyze subjects' behavior in the 9 subject-specific rounds of Part II.[25] In these 9 rounds, subjects are faced with choices that are determined by their decisions in Part I. For example, if in one of the rounds in Part I, a subject pays \$1 and gives \$16.30 to B, then it means that she prefers $(11.30, 16.30)$ to $(12.30, 12.30)$. In the corresponding Part II round, she faces a choice between those two allocations: $(11.30, 16.30)$ and $(12.30, 12.30)$. More generally, if in Part I she pays p and choose x_B^* for B over the $(12.30, 12.30)$ allocation, then in Part II she has to choose between the following two payoff pairs: $(12.30 - p, x_B^*)$ and $(12.30, 12.30)$. In other words, in Part II, subjects face a choice between the allocation they chose in Part I and an alternative allocation they could have chosen but didn't.

If in Part II A retains her choice of $(12.30 - p, x_B^*)$ over $(12.30, 12.30)$, then we say that she pays p in Part II to implement her desired allocation from Part I. Note that while the $(12.30 - p, x_B^*)$ allocation is identical to what A chose in Part I, in Part II paying p does not lead to additional power, it just leads to implementing this specific payoff distribution. Indeed, whether or not A pays p in Part II, the choice she faces is between two payoff pairs that are fixed, whereas in Part I, paying p allowed A to choose a precise

[25]The remaining 11 rounds are identical for all subjects and analyzed in Section 4.3.

payoff for B. If a subject's preferences are on distributional outcomes only, she should choose should be the same allocation in Part II as in Part I: the subject should choose $(12.30 - p, x_B^*)$.

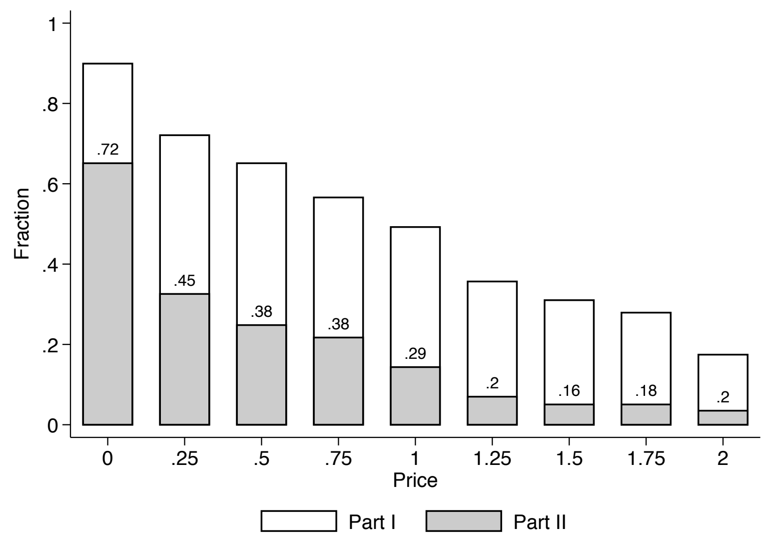

FIGURE 3. Fraction of subjects who pay in Part I and Part II, by price in USD.

Figure 3 adds shaded bars to Figure 1 that represent the fraction of subjects who pay in Part II conditional on paying in Part I.[26] The conditional fraction of subjects paying in Part II for each price is written at the top of each shaded area. For example, only 29.1% of subjects who pay a price of $1 in Part I also pay in Part II, i.e., they choose the same allocations as in Part I. There is a stark difference between subjects' willingness to pay in Parts I and II at the aggregate level for each price. For every p in $\mathcal{P}$, conditional on having paid in Part I, far fewer subjects in Part II are willing to pay the same price to ensure the payoff distribution they chose in Part I. Interestingly, there is a difference at the price of zero, which means that for some subjects, $(12.30, x_B^*) \succ (12.30, 12.30)$ when x_B^* is chosen by the subjects themselves but $(12.30, x_B^*) \prec (12.30, 12.30)$ when the same value of x_B^* is fixed. Thus, even at no cost to themselves, they do not implement their Part I allocations in Part II.[27]

[26]Recall from Section 2.2 that which payoff pair appeared on the left of the screen and which appeared on the right was randomly determined by the computer. So there is no bias coming from the $(12.30, 12.30)$ allocation being constantly on the same side of the screen.

[27]It is possible that a subject in Part I pays because she is indifferent between the two choices, and any choice reversal in Part II is the result of this indifference. If this were the case, and if subjects were only interested in final outcomes (as opposed to power), we would expect 50% of reversals in Part II. A two-sided test of proportion shows that the probability of paying in Part II conditional on having paid in Part I is statistically different than 50% with a p-value less than 0.01. This is also true if one looks at each price individually, with the exception of $p = 0.25$. In Section 4.2, we use the individual level data to provide further evidence that these switches are not due to indifference.

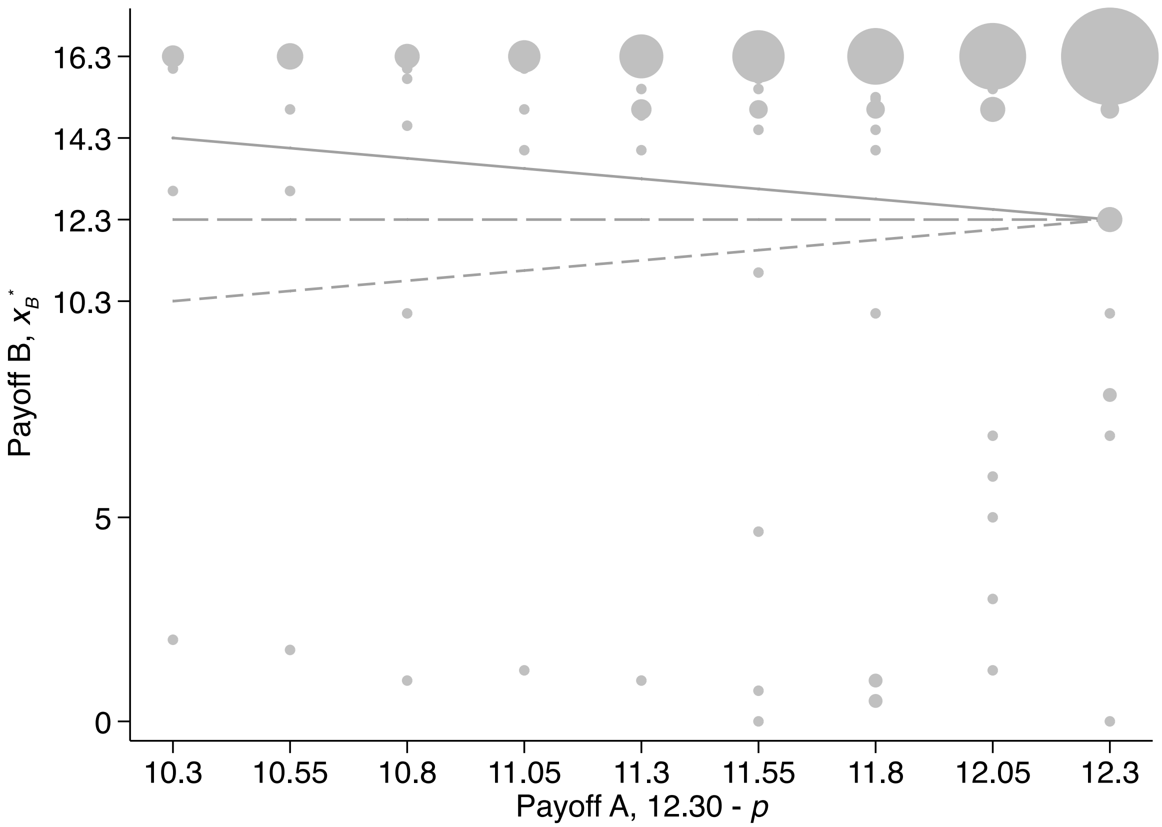

FIGURE 4. Payoff allocations $(12.30 - p, x_B)$ preserved by subjects in Part II.

Figure 4 shows the Part I $(12.30 - p, x_B^*)$ allocations that subjects preserved in Part II. Recall that in Part I, subjects decide to pay p in 1,148 cases. In Part II however, they choose not to pay in 686 or 59.8% of those cases, that is, they revert to the $(12.30, 12.30)$ allocation. Subjects revert 43.5% and 92.0% of efficient and competitive allocations, respectively.

Our aggregate results provide strong evidence that preferences for power are non-trivial. Many subjects are willing to pay if paying increases their power over the payoffs of others as is the case in Part I. However, they are much less willing to pay to implement the same payoff allocations when paying does not lead to additional power as is the case in Part II. If subjects' decisions to obtain the right to choose payoffs of others in Part I were driven entirely by their social preferences then there should be no reversals in Part II. Thus, our results suggest that (1) preferences for power exist and are substantial and (2) that they are different than and cannot be explained by social preferences. In the next section, we continue our analysis at the individual level and explore the broad categories of preference classes among our subjects.

4.2. **Preferences for Power: the Individual Level.**

4.2.1. *Definition.* We say that someone has preferences for power if she exhibits different behaviors in Part I and Part II. Specifically, someone has preferences for power if her willingness to pay is different in Part II than in Part I, i.e., $\bar{p}_{\mathrm{I}} \neq \bar{p}_{\mathrm{II}}$ (see Section 3). Following Assumption 4 of the theory, we assume that the subjects' choices at the maximum price they pay in each Part are the result of strict preferences.[28]

4.2.2. *Demand functions.* In the main text we focus on those subjects who have well-behaved demand functions in Parts I and II. Specifically, we require that in Part I there exists $\bar{p}_{\mathrm{I}}$ such that for all $p \leq \bar{p}_{\mathrm{I}}$ a subject pays p to choose B's payoff, and for all $p > \bar{p}_{\mathrm{I}}$ a subject does not. In addition, we require that in Part II there exists $\bar{p}_{\mathrm{II}}$ such that for all $p \leq \bar{p}_{\mathrm{II}}$ a subject chooses $(12.30 - p, x_B)$ and for all $p > \bar{p}_{\mathrm{II}}$ she does not.[29] Thus, we focus our analysis on subjects who have a single switching point in each Part, i.e. have step demand functions. Figure 5 graphically visualizes these well-behaved demand functions. Note that in Figure 5, $\bar{p}_{\mathrm{II}} < \bar{p}_{\mathrm{I}}$ but we include subjects for who $\bar{p}_{\mathrm{II}} < \bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}} = \bar{p}_{\mathrm{I}}$ in our analyses.

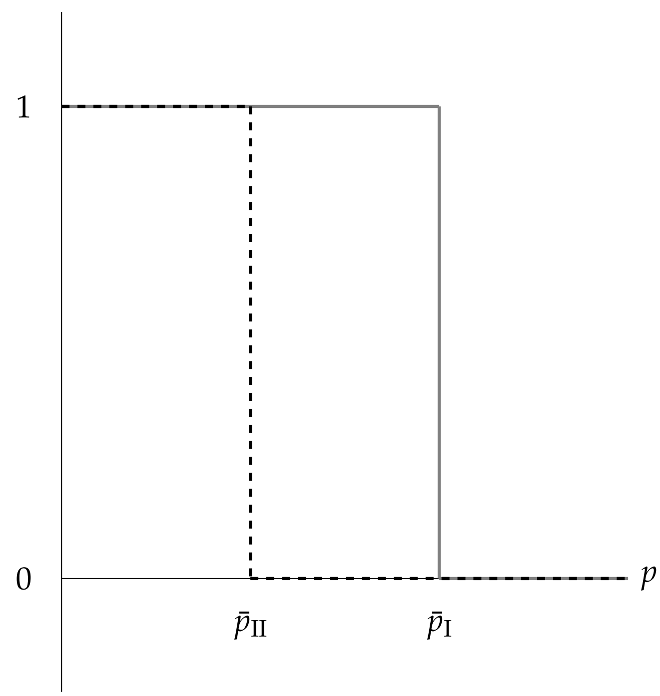

FIGURE 5. Well-behaved demand functions in Part I and Part II.

We have 126 subjects (48.8% of our sample) who have well-behaved demand functions in both Parts I and II. This proportion is relatively large given that prices are randomly

[28]We return to this point in Section 4.2.3 and show that our results are robust to relaxing this assumption.
[29]x_B is equal to x_B^*, her choice for B's payoff, if A paid p in Part I, and is equal to $12.30 + 2p$ if she did not.

drawn from the set $\mathcal{P}$ in every round in Part I and that the 9 rounds that correspond to Part I are randomly presented in Part II among the 20 Part II rounds.[30] There are no significant differences, either in terms of magnitude or statistically, in the distribution of skips in both Parts of the Power Game. For example, the fraction of subjects who make no skips in Part I (Part II) is 67.8% (67.8%, identically), who make one skip is 15.1% (11.6%), and who make two skips or more is 17.1% (20.5%). Thus, subjects understand Parts I and II equally well.

Note that by design our restricted sample removes the possibility that our results are due to time trends. Indeed, a subject who pays only in early rounds is unlikely to have a step-demand function since in Part I prices are randomly drawn from the set $\mathcal{P}$, and in Part II the order of rounds is random. In Appendices C and D we expand our sample to individuals who make one and any number of skips and show that our results are robust to these changes.

4.2.3. *Difference in willingness to pay across parts: preference classes.* Figure 6 shows the joint distribution of the subjects' willingnesses to pay in Parts I and II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, for those subjects who pay at least once. For example, for 3.97% of our subjects (5 out of 126 subjects), $\bar{p}_{\mathrm{I}} = 0.25$ and $\bar{p}_{\mathrm{II}} = 0.25$, while for 11.1% of our subjects (14 out of 126 subjects) $\bar{p}_{\mathrm{I}} = 2$ and $\bar{p}_{\mathrm{II}} = 0$. Subjects who are willing to pay more (less) in Part I than in Part II of the Power Game, i.e., for whom $\bar{p}_{\mathrm{I}} > \bar{p}_{\mathrm{II}}$ ($\bar{p}_{\mathrm{I}} < \bar{p}_{\mathrm{II}}$), appear below (above) the 45-degree line. Subjects whose willingness to pay is the same across Parts lie on the 45-degree line. Note that the joint distribution of the willingnesses to pay in Parts I and II is not concentrated nor symmetric around the 45-degree line. This is strong evidence that our results are not due to mistakes, indifference or confusion.

We use our theoretical predictions (see Table 2 in Section 3) to sort our subjects into different preference classes. Recall that these are only based on their willingnesses to pay in Parts I and II of the Power Game: different preference classes correspond to different relationships between $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$.

Subjects with standard preferences only care about their own payoff. These subjects are never willing to decrease it to affect the payoff of others. Thus, for them $\bar{p}_{\mathrm{I}} = 0$ and

[30] A total of 175 subjects (67.8% of our sample) skip one price or less across both Parts. For example, in Part I a subject may have a well-behaved demand function, but in Part II she pays all the way up to \$1, and then never pays for prices greater than \$1 except for \$1.50. Since such a subject has "skipped" the price of \$1.25, we call that a single skip across both Parts. If that subject instead does not pay for \$1.25 or \$1.50 but does for \$1.75, we call that two skips.

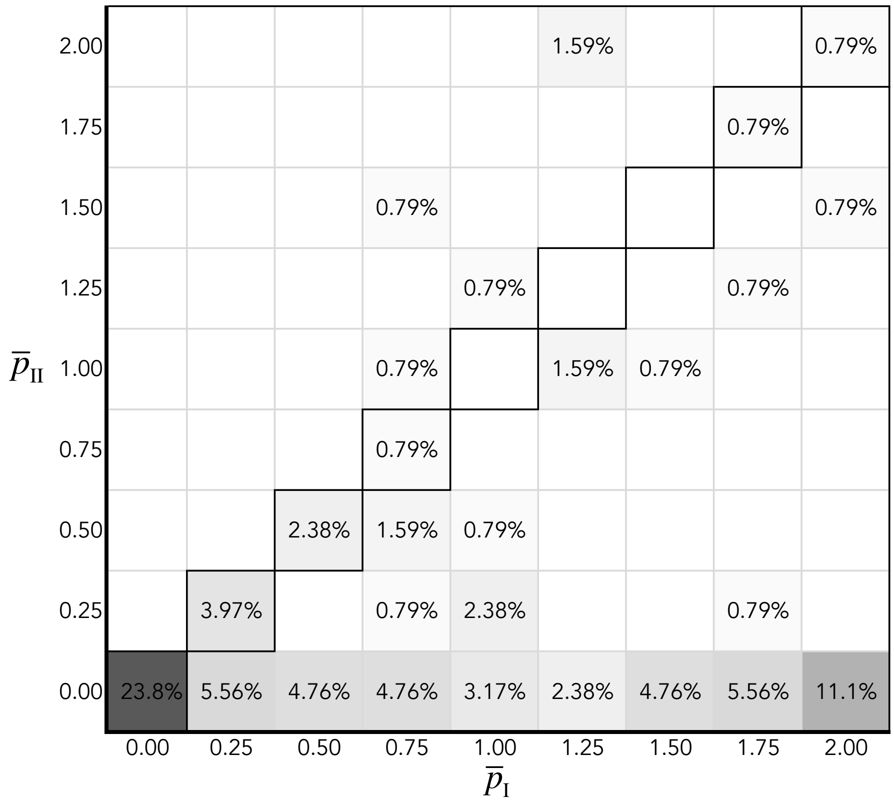

FIGURE 6. Joint distribution of the willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$.

$\bar{p}_{\mathrm{II}} = 0$.[31] We see from Figure 6 that 23.8% (30 out of 126) of our subjects have standard preferences.[32]

Recall that in our notation, the utility function takes three arguments, x_A, x_B, and λ. The last argument λ is equal to 1 if player A can choose a specific payoff of player B, i.e., pays to pick x_B from an interval. Otherwise, λ is 0. Thus, for subjects with power preferences and no social preferences $U(x_A, x_B, 1) > U(x_A, x_B^{'}, 0)$ for all $x_B, x_B^{'}$. In Figure 6 these subjects are located along the horizontal axis and together represent 42.1% of our sample (or 53 out of 126 subjects). Indeed, in Part I they are willing to pay up to $\bar{p}_{\mathrm{I}} > 0$ to choose the payoff of B, but in Part II never pay to implement the allocations they chose in Part I and choose to maximize their own payoff instead. As we can see from Figure 6, these subjects' willingnesses to pay in Part I spans the entire range of prices, from \$0.25 to \$2. In fact, 11.1% (or 14 subjects) are willing to pay up to the maximum price.

[31] We acknowledge here that the discreteness of $\mathcal{P}$ means that among these subjects some might have paid to choose others' payoffs, but the lowest positive price (\$0.25) is still too high for them.

[32] In principle, these subjects might instead have strong preferences for equality. If subjects only pay at a price of zero because they have preferences for equality, then they should consistently give \$12.30 both in that round as well as in round 10 of Part I, where they are forced to choose B's payoff. We find that none of these 30 subjects do so.

Subjects who have the same willingness to pay across both Parts of the Power Game, i.e. subjects for who $\bar{p}_{\mathrm{I}} > 0$ and $\bar{p}_{\mathrm{I}} = \bar{p}_{\mathrm{II}}$, have social preferences and no power preferences. They derive no additional utility from power but instead care about payoff distributions, independently of how they are attained. In particular, whether they chose x_B from an interval or not does not affect their utility: $\mathrm{U}(x_A, x_B, 1) = \mathrm{U}(x_A, x_B, 0)$. These subjects lie on the 45-degree line in Figure 6 and represent 8.7% of our sample (11 out of 126 subjects).

Subjects who have positive but different willingnesses to pay across Parts I and II of the Power Game have preferences for power and social preferences. For 10.3% of our subjects (or 13 or them) $\bar{p}_{\mathrm{I}} > \bar{p}_{\mathrm{II}} > 0$. These subjects clearly have social preferences since they pay positive prices in Part II. However, they are unwilling to pay up to p_{I} because in Part II paying does not lead to additional power. In other words, in Part I these subjects derive utility from the act of choosing a specific amount for B, as well as from the resulting distribution itself. In Part II however, they can only derive utility from the resulting distribution and so are willing to pay less. There are also subjects who pay more in Part II than in Part I: $\bar{p}_{\mathrm{II}} > \bar{p}_{\mathrm{I}} > 0$. These subjects have social preferences and dislike power: $\mathrm{U}(x_A, x_B, 0) > \mathrm{U}(x_A, x_B, 1)$. Only 4.0% (or 5 subjects) of our sample have this type of preference.

Finally, 11.1% of our sample (or 14 subjects) never pay in Part I, even at a price of zero. These can fall into several preference classes: they may have social preference but dislike power more than they care about others; they may like equality and dislike power or be indifferent to it; they may have standard preferences and dislike power or be indifferent to it. Since we cannot say with certainty how these subjects treat others or what their preferences for power are, we group them into an "unclassified" category.[33] This implies that the fraction of subjects with standard preferences represent a minimum of 23.8% of our sample, but could be as high as 34.9%. Similarly, subjects who have social preferences and dislike power are a minimum of 4.0% of our sample, but this fraction could be as high as 15.1%.

In Figure 7 we show the distribution of preference classes among our subjects. The most common class are preferences for power without social preferences, the Power+ preference class. These subjects represent 42.1% of our sample. The second largest class is the Standard preference class: these subjects neither care about power nor about others. They represent 23.8% of our sample. Together, these two categories comprise about two thirds of the sample. Grouping together subjects who have social preferences in any

[33]We are able to say that subjects with strong preferences for equality can be at most 3.2% (4 subjects) since this is the fraction of subjects who give \$12.30 in round 10 of Part I. In addition, we can rule out the possibility that some have preferences for efficiency and dislike power because out of these subjects, none pay in Part II.

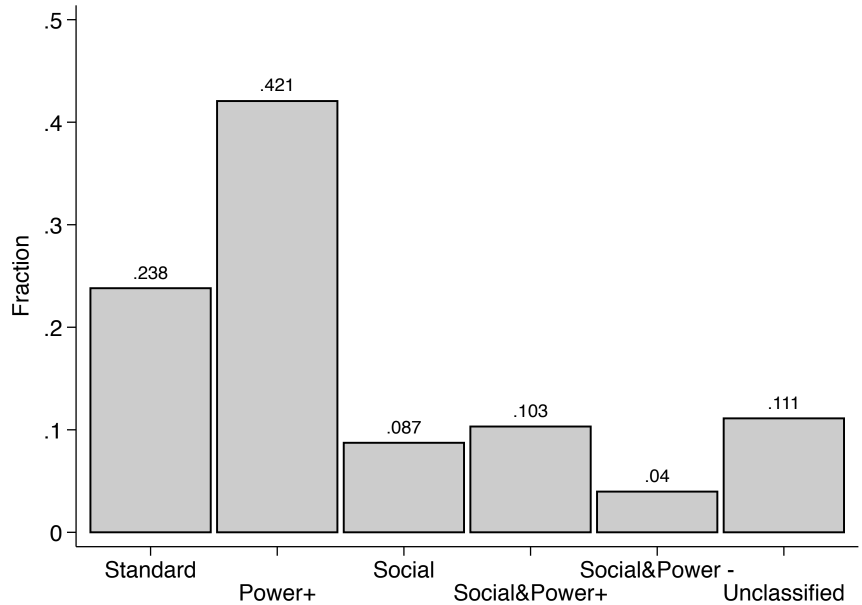

FIGURE 7. Distribution of preference classes.

capacity (the Social, Social&Power+ and Social&Power- classes) leads to roughly 23.0% of our subjects. Subjects who have preferences for power in any capacity (the Power+, Social&Power+ and Social&Power- classes) represent about 56.4% of our subjects.[34]

Regarding how much subjects are willing to pay in order to implement their preferences, we can glean from Figure 6 that subjects in the Power+ preference class are willing to pay the most, on average \$1.25, or over 10% of their potential payoff. In fact, more than half of our Power+ subjects are willing to pay \$1 or more in Part I. On average, subjects in the Social preference class are willing to pay \$0.66 to implement their preferences, while those in the Social&Power+ and Social&Power- classes are willing to pay \$1.21 and \$1.00, respectively.[35]

[34]Here we have assumed that subjects' choices reflect their strict preferences. Given the coarseness of $\mathcal{P}$, it is unlikely that indifference is attained at one of the discrete price points subjects face. However, here we still consider whether subjects' indifference at p_{I} or p_{II} might affect our results. For example, consider a subject with "true" willingnesses to pay of \$1.20 and \$1.00 in Parts I and II, respectively. Then, in Part I she would pay up to \$1.00, not pay at \$1.25, and for her $p_{\mathrm{I}} = \$1.00$. Meanwhile, in Part II she might pay or not pay at \$1.00, since she is indifferent at that price and therefore p_{II} may be either \$0.75 or \$1.00. Thus, this subject's indifference may lead to her being classified as belonging to either Social&Power+ preference class or to the Social preference class. As such, some subjects for who $|p_{\mathrm{I}} - p_{\mathrm{II}}| \leq 0.25$ may be mis-classified. However, since decisions made by subjects for who $|p_{\mathrm{I}} - p_{\mathrm{II}}| > 0.25$ cannot be explained by indifference, the Power+ preference class still would remain significant at 36.5%.

[35]This provides additional evidence that choices are deliberate. Indeed, if, for example, most of these subjects were willing to pay only up to 25 or 50 cents in Part I, one may have wondered if these were unintentional decisions. However, as is evident from Figure 6, the distribution subjects' willingnesses to pay is not such.

4.3. **Preference Classes and Predicted Behavior in Other Dimensions.**

Our preference classification depends only on the difference in subjects' willingnesses to pay across Parts I and II of the Power Game. If the classification indeed captures differences in preferences across subjects, then the identified preference classes should predict subjects' choices in other dimensions. Here we provide evidence that is indeed the case. First, subjects we have classified as having social preferences, regardless of their attitude towards power, are consistent in the amounts they give to type B players. In contrast, subjects in the Power+ preference class exhibit much more variation in their giving behavior both within and across subjects. Second, we show that these classes also predict subjects' decisions in tasks that are unrelated to Part I of the Power Game. More specifically, in the absence of power, subjects in the Power+ class behave much like subjects with standard preferences, i.e., they maximize their own payoff, while those with social preferences do not.

4.3.1. *Choice for player B's payoff.* In this section, we compare subjects identified as having social preferences and power preferences in terms of their giving behavior in rounds 1 through 9 in Part I of the Power Game. Subjects with social preferences belong to the following preference classes: Social, Social&Power+, Social&Power-. In other words, these subjects may be indifferent towards, like, dislike power, but they all have preferences towards B's payoff. In contrast, subjects in the Power+ preference class are indifferent towards the payoffs of others, i.e., they do not have social preferences. If subjects in the Power+ preference class are correctly identified, then we should see them behaving differently in terms of how they give to B compared with subjects who have social preferences in any capacity.

Figure 8 shows the cumulative distribution function of the amounts given to player B, averaged per subject, separately for subjects in the Power+ preference class and those who have social preferences. Figure 8 shows that there are clear differences between subjects in terms of what they choose to give to B and that these differences are consistent with their preference class. What is visually different is also different statistically.[36]

In addition, subjects with social preferences are very homogeneous: 89.7% of them always give the maximum allowable amount of \$16.30, and 96.6% give always give more to B than they themselves receive. In contrast, subjects in the Power+ preference class are very heterogeneous in terms of what they give to B and these amounts span almost the entire choice space, that is, the $[\$0, \$16.30]$ interval. Further, at the individual

[36]Kolmogorov-Smirnov and Wilcoxon-Mann-Whitney tests show that the distribution of amounts given by subjects in the Power+ preference class is statistically different than that of those who have social preferences, with p-values of less than 0.001. The unit of observation is the average amount given by each subject.

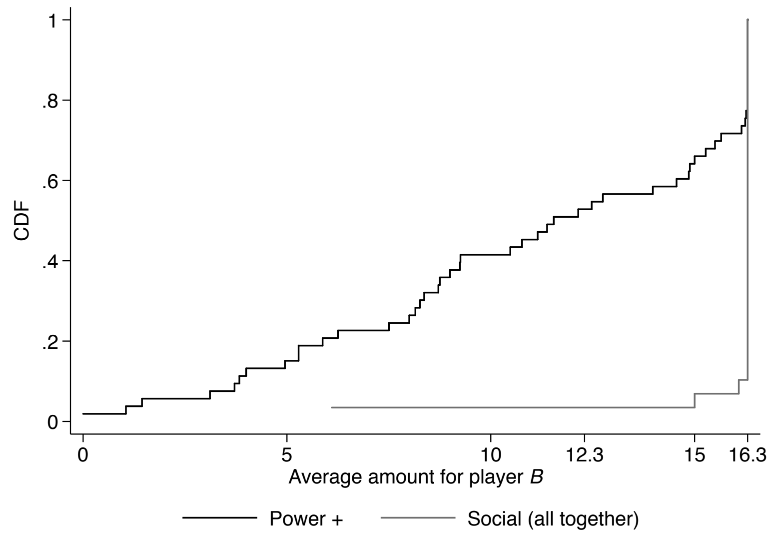

FIGURE 8. Distribution of the average amount given to player B.

level, among Power+ subjects there is much more variation within each subject's choice compared with subjects who have social preferences: the within-subject mean standard deviation of choices for B's payoff is higher for subjects in the former category than in the latter (1.63 versus 0.15).[37] In other words, it is not the case that Power+ subjects have strong but very different preferences towards B. Instead, they seem to have very weak preferences regarding B's payoff, as our classification implies.

We also compare subjects' giving behavior in the round in which price is zero with behavior in round 10 of Part I where all subjects are forced to choose for B. In both of these rounds, A players receive \$12.30. If subjects have social preferences, we expect them to give the same amount to type B players in these two rounds since their own payoff is identical in both cases. Among those subjects that we classify as having social preferences, 93.1% give the same amount in those two rounds while only 67.9% of Power+ subjects do so.[38]

The above results indicate that subjects in the Power+ preference class seem indeed to attach little importance to what B players earn. Their apparent indifference towards the payoffs of others, displayed both at the aggregate and individual levels, is in sharp contrast with the consistent giving behavior of those identified as having social preferences. These systematic differences in giving behavior between identified preference classes provide convincing evidence that the Power Game yields a meaningful preference classification.

[37] A two-sided test comparing these standard deviations shows that they are statistically different with a p-value less than 0.001.

[38] A two-sided test of proportions shows that these are statistically different with a p-value of 0.015.

4.3.2. *Behavior in a separate task.* Recall that in Part II of the Power Game, we present our subjects with 11 decision problems that are unrelated to their choices in Part I. Six of those problems (CR1 through CR6) are inspired by Charness and Rabin (2002).[39] In Table 5 we compare our subjects to those of Charness and Rabin (2002) and Chen and Li (2009) by presenting the proportion of subjects who choose the first option in each decision problem. In this table, the way we present the decision problems is such that the first option always yields a higher payoff for A players, except in problem CR1 where payoffs for A are identical across the two options. In the experiment, which option is presented on the left or on the right side of the screen is randomly and independently determined for each subject and for each decision problem. Our sample is largely similar to those in other institutions.[40] If anything, our subjects seem to choose the payoff-maximizing option more often than in Charness and Rabin (2002) and Chen and Li (2009).

TABLE 5. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task across three samples: Charness and Rabin (2002), Chen and Li (2009), and our sample.

Decision	First Option	Second Option	CR2002	CL2009	Our Subjects
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$	31%	33%	31%
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$	51%	82%	56%
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$	100%	NA	98%
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$	67%	76%	85%
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$	27%	50%	68%
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$	78%	64%	82%

In Table 6 we split our sample according to our subjects' preference classes. We present the fraction of subjects who choose the first option in decision problems CR1-CR6 as well as the 5 remaining independent choice problems (PT1-PT5) for subjects with different preferences: Standard, Power+, and social preferences.[41]

Importantly, in all Part II rounds, in each of these decision problems subjects cannot increase their power by sacrificing some of their payoff, since the payoff for B is fixed in both options. That is, the amount of power subjects have is the same irrespective of which option they choose. Thus, any difference in behavior across subjects with different preference classes can only be due to their preferences beyond those for power.

[39]They are re-scaled so that the numbers are similar to those our subjects face in Part I.

[40]Here we present the results only for those subjects who have well-behaved demand functions in Parts I and II of the Power Game. In Appendices C and D we expand our sample to individuals who make one or any number of skips and show that none of our results change.

[41]This last category includes Social, Social&Power+, Social&Power- subjects.

When additional power is not attainable, our theory (see section 3) predicts that individuals in the Power+ preference class should behave similarly to individuals in the Standard preference class. That is, subjects with Standard and Power+ preferences should be equally likely to choose the payoff-maximizing option. In contrast, subjects with social preferences should not choose the first option more often than subjects with no social preferences. Note that subjects with social preferences do not necessarily always choose the second option since their choices depend on each subject's marginal rates of substitution between their own and B's payoffs. The last two columns of Table 6 are populated with check marks and crosses. A check mark indicates that a test of proportion comparing specific groups is consistent with the theory. Specifically, comparing the behavior of subjects in the Power+ and Standard preference classes, for all decision problems but CR1, a check means that we cannot reject the null that the two proportions are equal in favor of the alternative that they are different, or in other words, that the two-sided p-values are greater than 10%. For CR1, since A's payoff is identical across both options, any proportion is consistent with the theory. When comparing Power+ versus social preferences, a check means that we cannot reject the null in favor of the latter choosing the first option in greater proportion, or in other words, that the one-sided p-values are greater than 10%.

Table 6. Fraction of subjects choosing the first option in the independent tasks by preference class.

	Option		Preference Class			Consistent w/ Theory	
Decision	First	Second	Standard	Power +	Social (All)	PP vs. StP	PP vs. SP
CR1	(6.60, 6.60)	(6.60, 12.30)	13%	49%	7%	✓	✓
CR2	(6.60, 6.60)	(6.20, 12.30)	63%	74%	7%	✓	✓
CR3	(3.10, 12.30)	(0.00, 0.00)	97%	98%	100%	✓	✓
CR4	(10.50, 5.30)	(8.80, 12.30)	97%	85%	76%	✗	✓
CR5	(12.30, 3.50)	(10.50, 10.50)	80%	72%	41%	✓	✓
CR6	(12.30, 0.00)	(6.15, 6.15)	90%	79%	79%	✓	✓
PT1	(10.10, 5.20)	(9.10, 9.10)	87%	70%	52%	✗	✓
PT2	(12.30, 5.10)	(10.10, 12.30)	90%	81%	66%	✓	✓
PT3	(12.30, 9.60)	(9.60, 12.30)	100%	98%	100%	✓	✓
PT4	(12.30, 7.80)	(7.80, 5.40)	100%	96%	100%	✓	✓
PT5	(6.15, 6.15)	(0.00, 0.00)	100%	100%	100%	✓	✓

As is clear, across almost all decisions problems, subjects make choices that are consistent with our preference classification and our theory. The last column shows that the fraction of subjects with social preferences choosing the first option is never greater than that of Power+ subjects, in line with the theory. In fact, the latter fraction is statistically greater than the former in 5 of the 6 decision problems in which the second option is more efficient than the first, as the theory and our previous analyses predict.[42] According to the penultimate column, Power+ and Standard subjects with behave similarly in most cases. There are two exceptions, CR4 and PT1, where we find that the fraction of subjects choosing the first option in the two groups is different, though the statistical significance is marginal with p-values at 0.098 and 0.085, respectively.

Aggregating behavior for each subject across all decision problems, we find that the fraction of subjects who always choose the payoff-maximizing option among those with standard preferences is 53.3%. For Power+ subjects this fraction is 47.2% and it is 6.9% for those who have social preferences. The fractions for the Power+ and Standard preference classes are not statistically different, while the fraction for subjects who have social preferences is significantly smaller than either of the two other categories.[43]

Finally, we note that in PT3-PT5, almost all subjects choose the first option, irrespective of their preference class. These decision problems are included in our design to test whether subjects understand our game and instructions. For example, PT3 serves a specific role as it allows us to show that subjects understand that they are to act as type A players. This is the case since if they had any doubts more subjects would have chosen the second option. PT4 and PT5 are chosen to make sure that subjects care about payoffs. Overall, our results in PT3-PT5 demonstrate that subjects with different preferences understand our experiment equally well. Thus, the differences in subjects' behavior across preference classes cannot be explained by misunderstanding or confusion regarding roles or payments.

4.4. **Discussion.**

Here we open a discussion on whether behavior that we identify as preferences for power may in fact be due to other factors. Since subjects who have preferences for power are those who pay in Part I and not in Part II (or who pay differently in Part II if we consider those subjects who also have social preferences), we must consider whether other factors independent of preferences could lead to such choices.

4.4.1. *Uncertainty regarding type assignment.* Recall that in our experimental implementation, subjects are not informed of their true type. Subjects are asked to make decisions as

[42]These are CR1, CR2, CR5, PT1 and PT2. In these problems we can reject the null that the two fractions are equal in favor of the alternative that the fraction for Power+ subjects is strictly greater than that for subjects with social preferences. The p-values range from less than 0.001 to 0.058.

[43]The p-value for a two-sided test of proportions comparing the first two is 0.589, while the p-values for each pairwise comparison between social and the others are smaller than 0.001.

if they were type A, since if their true type turns out to be B, none of their decisions matter for their payment. Subjects' uncertainty about their type might create two potential problems. The first is that it might generate confusion and lead subjects to incorrectly believe that their decisions might matter for their payoff even if their true type turns out to be B. Such a subject would incorrectly believe that if she pays she receives \$12.30 $- p$ *or* \$16.30, depending on her realized type, and receives \$12.30 for sure if she doesn't pay.[44] The second potential problem is that not knowing what type one is might make a subject more likely to empathize with a type B player, and possibly exacerbate her social preferences. However, if either of these two factors affects a subject's decisions in Part I, it should also affect her decisions in Part II in the same way. In other words, these subjects' paying behavior should be identical across parts. Thus, type uncertainty has no impact on our identification of power preferences, but might lead us to over-identify social preferences relative to Standard preferences compared with a design with no type uncertainty.

4.4.2. *Differences in the number of stages across Parts.* Here we consider whether the fact that Part I has two stages while Part II has a single stage may explain our results. Indeed, in Part I, subjects must first decide whether to pay or not and only then choose how much to give to player B. If subjects are confused by the additional stage, they may pay more often in Part I than in Part II. To assess this, we take our entire sample and look at the distribution of skips in Parts I and II. We see that they are no different. Remember that a subject has well-behaved demand functions in Parts I and II if in each of those Parts she pays up to a certain price and then switches to not paying. If a subject makes one skip, e.g., pays up to \$1.00, does not pay at \$1.25, pays at \$1.50, and never pays afterwards, we say that she makes one skip. If, instead, she does not pay at \$1.25 and \$1.50 but pays at \$1.75 and then never again, we call that two skips, etc.

The fraction of subjects who make one skip in Part I (Part II) is 15.1% (11.6%), who make two skips or more in Part I (Part II) is 17.1% (20.5%). A Wilcoxon matched-pairs signed-rank test (Wilcoxon (1945)) shows that subjects are not more likely to skip prices in Part I than in Part II.[45] Further, the fractions of subjects who have well-behaved demand functions in Part I and II are exactly the same, 67.8%. Thus, there are no differences in the subjects' understanding of Parts I and II of the Power Game that might affect our preferences classification.

[44] If such a subject pays, she would choose to give herself \$16.30, the maximum allowable amount. How much she weighs the chances of receiving \$12.30 $- p$ versus \$16.30 would depend on what she believes the probability of being a type A player is. If she doesn't pay, both A and B receive \$12.30.

[45] The p-value on that test is greater than 10%.

4.4.3. *Mistakes explain preference classes.* While there are no differences in the number of skips across the two Parts of the Power Game for all subjects, one might argue that price skips could be concentrated within subjects in the Power+ class. In other words, one might say that what we mis-identify preferences, in particular those for power, and that these are simply due to noise and randomness in subjects' paying behavior. This point is moot in the main text since in our main text we focus only on those subjects who have well-behaved demand functions and make no skips. However, this issue might exist if we consider all subjects and allow any number of skips. Here we address this potential concern.[46]

In the entire sample, we use two different methods to determine subjects' willingness to pay: *local* and *global maximum* methods. In the *local maximum* method, a subject's willingness to pay is defined as the maximum price p at which she pays before making her first skip, or 0 if she does not pay at a price of 0. In the *global maximum* method, a subject's willingness to pay is defined as the global maximum price she pays, or 0 if a subject never pays at any price. Which preference class a subject fits into depends on which method is used. For example, suppose that in Part I a subject pays all prices until \$1.00 and then pays once more at \$1.75, i.e., makes two skips, at \$1.25 and \$1.50. Suppose further that in Part II she pays only at a price of \$1.75, i.e., makes seven skips, one at each price up to \$1.50. This subject thus makes a total of 9 skips. According to our local maximum method, this subject's willingnesses to pay in Part I and Part II are \$1.00 and \$0, respectively, and she has Power+ preferences, since $p_I > 0$ and $p_{\text{II}} = 0$. According to our global maximum method however, this subject's willingnesses to pay in Part I and Part II are the same and equal \$1.75, and she Social preferences, since $p_I = p_{\text{II}}$.

In Panel A of Table 7 we report the distribution of total skips by preference class identified using the local maximum method. Using a series of pairwise Kolmogorov-Smirnov tests, we show here is no difference in the distribution of total skips between those classified as having standard preferences, those in the Power+ class and those who have social preferences.[47,48] In Panel B of Table 7, we report the distribution of Part I skips for preference classes defined using the global maximum method described above. Note that in this case, by construction, only subjects who never pay in both Parts of the Power Game are identified as having standard preferences and so for all these subjects the number of skips is zero. Also by construction, Power+ subjects make no skips in Part II. Thus, we report Part I skips only, and compare Power+ subjects with subjects

[46]The detailed analyses of our entire sample can be found in Appendix D.

[47]All p-values are strictly greater than 10%.

[48]In addition, using a series of pairwise Kolmogorov-Smirnov tests, we find that there is no difference in terms of distribution of skips in Part I or in Part II across our preference classes.

TABLE 7. Number of skips in the Power Game by preference class identified using the *local* and *global maximum* methods.

Preferences Class	Subjects	Mean	St.Dev.	Min.	p25	Median	p75	Max.
Panel A: *Local maximum* method - total skips								
Standard	59	1.98	2.81	0	0	0	3	12
Power +	104	1.52	2.34	0	0	0	2	9
Social (All)	80	1.60	1.91	0	0	1	2.5	9
Social	25	2.24	2.63	0	0	1	4	9
Social & Power +	41	1.17	1.22	0	0	1	2	5
Social & Power −	14	1.71	1.86	0	0	1	4	5
Panel B: *Global maximum* method - Part I skips								
Standard	30	0	0	0	0	0	0	0
Power +	87	0.77	1.23	0	0	0	1	5
Social (All)	126	0.88	1.48	0	0	0	1	7
Social	31	0.48	1.00	0	0	0	1	4
Social & Power +	55	1.22	1.71	0	0	1	2	7
Social & Power −	40	0.73	1.40	0	0	0	1	6

who have social preferences. We find no different in Part I skips across these preference classes.[49]

Thus, it is not the case that subjects in the Power+ preference class are simply those who make more mistakes than others.

4.4.4. *Time trends.* A fourth potential confound is that our results are simply due to time trends. This may be the case if, for example, individuals' power or social preferences can be satiated.[50] Two elements rule out this possibility. First, the instructions were very clear that only a single round in the experiment would be chosen for payment. Thus, actual power or generosity can only happen if a subject consistently implements her preferences in every round. Second, in our main text and analyses we focus on individuals who have well-behaved demand functions. This de facto controls for time. Indeed, if a subject were to decide to stop paying after a certain number of rounds (regardless of the reason why), it is very unlikely that her demand function would be well-behaved since prices are randomly drawn from the set $\mathcal{P}$.[51] Further, in Appendices C and D we use samples that allow for skips and our results are unchanged. Thus, our findings are not due to time trends.

[49]The p-value in a Kolmogorov-Smirnov test is greater than 10%.

[50]This may occur if after a few round subjects feel that they've exercised enough power or feel that they've done enough good deeds for the day.

[51]The same reasoning holds if subjects suddenly have an epiphany, for example suddenly realize that paying actually lowers their payoff.

4.4.5. *Warm glow.* Finally, one may also wonder whether elements such as "warm glow" could in fact explain our results. In Andreoni (1990), the author defines warm glow in a public good context and shows that an individual may contribute to a public good not because she cares about the public good per se, but because giving makes her feel good about herself. This brings about the possibility that we mis-identify our subjects' motives when making decisions. In Andreoni (1990) or papers that test his theory,[52] warm glow is not related to the size of the set of alternatives. In our experiment, this means that subjects who experience warm glow would experience the same level of it in both Parts of the Power Game. Thus, strictly speaking, an individual who makes decisions because she is motivated by warm glow should make the same decisions in both Parts of the Power Game. In our classification these subjects have social preferences. As such, they may in fact be motivated by warm glow.

Moving away from a strict interpretation of Andreoni's concept, one may wonder whether the level of warm glow increases with the size of the choice set. For example, someone paying in Part I who is generous to player B might experience warm glow because she knows there were many lower payoffs that she could have chosen for B but didn't. In Part II of our game the intensity of warm glow could be lessened by the fact that there are only two fixed alternatives for B's payoff. If this motivated the decision to pay in Part I and not pay in Part II, as our Power+ subjects do, we should see that they give higher amounts to B in Part I relative to our other subjects. However we observe quite the opposite: it is precisely those who pay in both parts who are generous. Those who pay positive prices in Parts I and II give an average of \$15.95, while those who only pay in Part I give far less at \$10.70. This difference is large both in magnitude and statistically.[53] In fact, more than half of the subjects who only pay in Part I , i.e. our Power+ subjects, give less than \$12.30 to B, which is what B would have received had A not paid. This fraction is only 3.5% among those subjects who pay in both Parts.

Thus, warm glow, whether in a strict sense or not, is not consistent with the behavior of our Power+ subjects.

5. Conclusion

In this paper we introduce a new game, the Power Game, and use it to identify individuals who have preferences for power without confounding other elements that may exist in the presence of power. Our work is the first to identify such preferences. We find that over half of the population values power *per se*, beyond its instrumental value.

[52]See, for example, Andreoni (1995), Crumpler and Grossman (2008).

[53]Both Kolmogorov-Smirnov and Wilcoxon-Mann-Whitney tests show that the distributions of amounts given to B are different with $p < 0.001$, where the unit of observation is the average amount given per subject.

We show that these preferences for power are different than, and cannot be explained by, social or other-regarding preferences.

In particular, the vast majority of subjects who value power, do so in the absence of social preferences: they attach little value to other people's outcomes and instead enjoy being the ones to choose those outcomes. Given that individuals with other regarding preferences tend to be generous rather than petty and spiteful, our results imply that social welfare is likely to decrease when individuals with power preferences obtain power.

Thus, understanding people's motives in seeking positions of power has important implications. This is the case in various settings, for example as it relates to the growth of a firm, the expansion of an organization or the direction a country takes. Our design itself can perhaps be used as a tool to identify those who have preferences for power per se, as opposed to those who might use power to implement their social preferences and increase social welfare.

Up until now, desires for autonomy, control and power had not been disentangled. Disentangling power preferences from other non-pecuniary motives marks an important step towards understanding why individuals may seek positions of power. We show that a large fraction of individuals seek power even if it does not grant them more control or autonomy from others. This suggests that while these may be present, a desire for autonomy and non-interference from others are not the only motivation to climbing the ladder to power. Consequently our findings provide strong reasons for incorporating non-pecuniary benefits in the design of contracts that relate to power.

Understanding people's preferences as they relate to others is a challenging task and great strides have been made towards gaining a better understanding of them. Up until now, preferences towards others, or the study of social preferences, has been the main focus of such work. We show that a substantial fraction of the population enjoys the *process* of choosing payoffs, rather than the resulting distribution itself, and is willing to give up substantial amounts of money in order to engage in this process. We hope that our work will serve as a catalyst for new empirical and theoretical research in this area.

References

Abbink, Klaus, and Abdolkarim Sadrieh. 2009. "The pleasure of being nasty." *Economics Letters*, 105(3): 306–308.

Aghion, Philippe, and Patrick Bolton. 1992. "An incomplete contracts approach to financial contracting." *The Review of Economic Studies*, 59(3): 473–494.

Andreoni, James. 1990. "Impure altruism and donations to public goods: A theory of warm-glow giving." *The Economic Journal*, 100(401): 464–477.

Andreoni, James. 1995. "Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments." *The Quarterly Journal of Economics*, 110(1): 1–21.

Bartling, Björn, and Urs Fischbacher. 2011. "Shifting the blame: On delegation and responsibility." *The Review of Economic Studies*, 79(1): 67–87.

Bartling, Björn, Ernst Fehr, and Holger Herz. 2014. "The intrinsic value of decision rights." *Econometrica*, 2005–2039.

Bolton, Gary E, and Axel Ockenfels. 2000. "ERC: A theory of equity, reciprocity, and competition." *The American Economic Review*, 166–193.

Brown, Alexander L, and Paul J Healy. 2016. "Separated decisions." *The Ohio State University working paper*.

Charness, Gary, and Matthew Rabin. 2002. "Understanding social preferences with simple tests." *The Quarterly Journal of Economics*, 117(3): 817–869.

Charness, Gary, David Masclet, and Marie Claire Villeval. 2014. "The dark side of competition for status." *Management Science*, 60(1): 38–55.

Chen, Yan, and Sherry Xin Li. 2009. "Group identity and social preferences." *The American Economic Review*, 99(1): 431–457.

Coffman, Lucas C. 2011. "Intermediation reduces punishment (and reward)." *American Economic Journal: Microeconomics*, 3(4): 77–106.

Cox, James C, Daniel Friedman, and Steven Gjerstad. 2007. "A tractable model of reciprocity and fairness." *Games and Economic Behavior*, 59(1): 17–45.

Cox, James C, Daniel Friedman, and Vjollca Sadiraj. 2008. "Revealed altruism." *Econometrica*, 76(1): 31–69.

Crumpler, Heidi, and Philip J Grossman. 2008. "An experimental test of warm glow giving." *Journal of Public Economics*, 92(5): 1011–1021.

Demsetz, Harold, and Kenneth Lehn. 1985. "The structure of corporate ownership: Causes and consequences." *Journal of Political Economy*, 93(6): 1155–1177.

Dessein, Wouter, and Richard Holden. 2017. "Organizations with power-hungry agents." *Working paper*.

Dyck, Alexander, and Luigi Zingales. 2004. "Private benefits of control: An international comparison." *The Journal of Finance*, 59(2): 537–600.

Falk, Armin, and Urs Fischbacher. 2006. "A theory of reciprocity." *Games and Economic Behavior*, 54(2): 293–315.

Fehr, Ernst, and Klaus M Schmidt. 1999. "A theory of fairness, competition, and cooperation." *The Quarterly Journal of Economics*, 114(3): 817–868.

Fehr, Ernst, Holger Herz, and Tom Wilkening. 2013. "The lure of authority: Motivation and incentive effects of power." *The American Economic Review*, 103(4): 1325–1359.

Fershtman, Chaim, and Uri Gneezy. 2001. "Strategic delegation: An experiment." *RAND Journal of Economics*, 352–368.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics*, 10(2): 171–178.

Grossman, Sanford J, and Oliver D Hart. 1986. "The costs and benefits of ownership: A theory of vertical and lateral integration." *Journal of Political Economy*, 94(4): 691–719.

Hart, Oliver D, and John Moore. 1995. "Debt and seniority: An analysis of the role of hard claims in constraining management." *American Economic Review*, 85(3): 567–585.

Hart, Oliver D, and John Moore. 2005. "On the design of hierarchies: Coordination versus specialization." *Journal of Political Economy*, 113(4): 675–702.

Levine, David K. 1998. "Modeling altruism and spitefulness in experiments." *Review of Economic Dynamics*, 1(3): 593–622.

Orhun, A Yesim. 2017. "Perceived Motives and Reciprocity." *Working paper*.

Owens, David, Zachary Grossman, and Ryan Fackler. 2014. "The control premium: A preference for payoff autonomy." *American Economic Journal: Microeconomics*, 6(4): 138–161.

Rabin, Matthew. 1993. "Incorporating fairness into game theory and economics." *The American Economic Review*, 1281–1302.

Wilcoxon, Frank. 1945. "Individual comparisons by ranking methods." *Biometrics Bulletin*, 1(6): 80–83.

Zizzo, Daniel John, and Andrew J Oswald. 2001. "Are people willing to pay to reduce others' incomes?" *Annales d'Economie et de Statistique*, 39–65.

Appendix A.

Below we present the instructions that the subjects received in our experiment.

INSTRUCTIONS

This experiment is in two parts. In each part you will participate in a number of Rounds. Only one part will be chosen for payment and only one Round in that part will count towards payment. Therefore, it is in your best interest to treat each Round independently and to treat it as if it were the one that mattered for payment. In addition to what you will earn in the experiment, you will get a 5-dollar participation fee if you complete the experiment.

Before we begin this experiment, you will be assigned a Type. You will be a Type A Player or a Type B Player. Your Type will remained fixed throughout this entire experiment.

At the start of each round, each Type A Player will be randomly rematched with a Type B Player. You will not know who you are matched with. In this experiment, only Type A players make decisions that matter for payment, and these decisions affect the payoff of both the Type A Player and the Type B Player he/she is matched with.

Even though your Type is determined at the start of the experiment and will remain fixed for the entire experiment, you will not know which Type of Player you are, until the end of the experiment. Since you do not know which Type of Player you are assigned to be, and since only Type A Players make decisions that matter for payment, we will ask everyone to make decisions as if they were Type A players. Your Type will be revealed to you only at the end of the experiment.

Please note that your Type will remain fixed throughout this entire experiment and at no point will you change roles. Your "true" Types have already been determined by the computer, and your decisions when acting as Player A **CANNOT** affect you or anyone else in this room if your "true" Type turns out to be Type B. In other words, if it turns out you are a Type B Player, no decision you make here can affect anyone's payoff, including your own. If it turns out your "true" Type is A, there is nothing that anyone else can do that will affect your payoff, and your decisions affect both your payoff and the payoff of the Type B Player you are matched with. Therefore, when making decisions, you should act as Player A. Further, since only "true" Type A Players make decisions that matter for payment in this experiment, in the remainder of the instructions we will assume you are a Type A Player.

Part I and Part II are different and independent of each other. We will give you the instructions for Part II of the experiment once you have completed Part I of the experiment. Below are the instructions for Part I of the experiment.

Part I

In this part of the experiment, you will make decisions over the course of 10 Rounds. As a Type A Player, in each Round, you can choose to pay a certain amount of money to obtain the right to choose the payoff of the Type B Player you are matched with. The price of that right will vary from Round to Round.

If you choose to pay this amount, you will be given \$12.30 and the price will be subtracted from the \$12.30 you have. If you obtain the right to choose the payoff for the Type B Player you are matched with, you can choose any number between 0 and \$16.30, both included, by increments of 5 cents.

If you choose to not pay that amount of money, you do not obtain the right to choose Type B's payoff. In this case, Player B will earn \$12.30 and you will choose your own payoff that can be any number between 0 and \$12.30, both included, by increments of 5 cents.

Example
Suppose in one of the Rounds the price of obtaining the right to choose the payoff for Player B is \$1.

- If you choose to pay \$1 then you can choose Player B's payoff between \$0 and \$16.30:
  - Suppose you choose \$3 as Player B's payoff. In this case, if this Round is chosen for payment, you will earn $\$12.30 - \$1 = \$11.30$ and the Type B Player you are matched with will earn \$3.
  - If instead you choose \$14.55 as Player B's payoff, and if this Round is chosen for payment, you will still earn $\$12.30 - \$1 = \$11.30$ and the Type B Player you are matched with will earn \$14.55.
- If you choose not to pay \$1 then you cannot choose Player B's payoff. Player B will earn \$12.30 and you will choose your own payoff between 0 and \$12.30:
  - Suppose you choose \$5.10 for yourself. In this case, if this Round is chosen for payment, you will earn \$5.10 and the Type B Player you are matched with will earn \$12.30.

- Suppose you choose \$10 for yourself. In this case you will earn \$10 and Player B will earn \$12.30.

You will play 10 Rounds of this game.

Remember that you will not change roles in this experiment. So as a Type A Player, your payoff will never be determined by someone else in this room. Also remember that only one Part of the experiment will be chosen to count for payment. If this Part is chosen to count, only **one** Round will matter for payment. So it is in your best interest to treat each Round as if it were the one that mattered for payment.

Before we start the 10 Rounds, I will show you two screens so that you can familiarize yourselves with the interface. The first screen will be what you would see if you **did** pay for the right to choose Type B's payoff. The second screen will be what you would see if you **did not** pay for the right to choose Type B's payoff. These "practice" screens do not count towards payment.

Part II

In each Round of this part of the experiment, you will be asked to choose between two options that will determine payoffs for both you and the Type B Player you are matched with. Here is an example of such a choice you can encounter in one of the Rounds (the choices you face will be different):

Your Payoff: \$12.30 Type B's Payoff: \$12.30	Your Payoff: \$9.15 Type B's Payoff: \$12.00

Here is an example, assuming that Part II and this Round was chosen for payment. If you choose the pair on the left, you will earn \$12.30 and the Type B Player you are matched with will earn \$12.30. If choose the pair on the right, you will earn \$9.15 and the Type B Player you are matched with will earn \$12.

You will play 20 Rounds of this game.

Appendix B.

Below we list the questions that we asked all subjects to answer at the end of the experiment.

I. Demographics
   (a) Please, enter your age:
   (b) Which gender do you identify with?
      - Male
      - Female
      - Other
   (c) How many years of university education have you received?
   (d) Are you a native English speaker?
   (e) Which faculty best describes your field of study?
      - Science
      - Social science
      - Arts
      - Engineering
      - Business
      - Other
   (f) What is the ZIP code of the place where you grew up?
   (g) What is the highest degree your mother has?
      - Less than high school
      - High school or equivalent
      - Some college
      - College
      - More than college
      - Other
      - Not sure
   (h) What is the highest degree your father has?
      - Less than high school
      - High school or equivalent
      - Some college
      - College
      - More than college
      - Other
      - Not sure
   (i) Have you ever participated in similar experiments before?

II. Understanding of the game

(a) Was anything confusing?
(b) What motivated your choices in this experiment?
(c) What do you think the experiment was about?

III. For the statements below, please indicate how they apply to you on a scale from 1 to 7:

1. I prefer a job where I have a lot of control over what I do and when I do it.
2. I prefer a job where I have a lot of control over what others do and when they do it, for example manager's job.
3. I enjoy political participation because I want to have as much of a say in running government as possible.
4. I try to avoid situations where someone else tells me what to do.
5. I would prefer to be a leader rather than a follower.
6. I enjoy being able to influence the actions of others.
7. I prefer a job where I have a say on promotion and pay of others.
8. I am careful to check everything on an automobile before I leave for a long trip.
9. Others usually know what is best for me.
10. I enjoy making my own decisions.
11. I enjoy having control over my own destiny.
12. I would rather someone else took over the leadership role when I'm involved in a group project.
13. I consider myself to be generally more capable of handling situations than others are.
14. I'd rather run my own business and make my own mistakes than listen to someone else's orders.
15. I like to get a good idea of what a job is all about before I begin.
16. When I see a problem I prefer to do something about it rather than sit by and let it continue.
17. When it comes to orders, I would rather give them than receive them.
18. I wish I could push many of life's daily decisions off on someone else.
19. When driving, I try to avoid putting myself in a situation where I could be hurt by someone else's mistake.
20. I prefer to avoid situations where someone else has to tell me what it is I should be doing.
21. There are many situations in which I would prefer only one choice rather than having to make a decision. I like to wait and see if someone else is going to solve a problem so that I don't have to be bothered by it.

Appendix C.

Here we redo all Figures and Tables that appear in Section 4.2 and Section 4.3 of the Main text, but allow for at most a single skip in the demand functions across the two Parts of the Power Game. For example, suppose that in Part I a subject pays all prices until \$1.00 and then pays once more at \$1.50, "skipping" the price of \$1.25. Suppose this subject has a well behaved demand function in Part II and pays for all prices up until \$1.00 and never after. We then say that this subject made a single skip across both parts of the Power Game. 126 subjects make no skips and an additional 48 subjects skip one price across both Parts.

In the entire sample, we use two different methods to determine subjects' willingness to pay: *local* and *global maximum* methods. In the *local maximum* method, a subject's willingness to pay is defined as the maximum price p at which she pays before making her first skip, or 0 if she does not pay at a price of 0. In the *global maximum* method, a subject's willingness to pay is defined as the global maximum price she pays, or 0 if a subject never pays at any price. Which preference class a subject fits into depends on which method is used. Take our example above, where a subject makes a single skip in Part I of the Power Game. According to our local maximum method, this subject's willingnesses to pay in Part I and Part II are identical and equal \$1.00, and she has Social Preferences, since $p_{\mathrm{I}} > 0$ and $p_{\mathrm{II}} = p_{\mathrm{I}}$. According to our global maximum method, however, this subject's willingnesses to pay in Part I and Part II are \$1.50 and \$1.00, respectively. She would therefore be classified as having Social&Power- preferences, since $p_{\mathrm{I}} > 0$ and $p_I \neq p_{\mathrm{II}}$.

For those subjects who make no skips we calculate their willingness to pay in the same way as in the Main text. That is, in each Part we use the maximum price at which they decide to pay before switching to not paying.

Figures C1 and C2 show the joint distribution of the willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *local maximum* and *global maximum* methods. These figures complement Figure 6 from the Main text. Figures C3 and C4 show the proportion of subjects with various preference classes based on their willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, where the preferences classes are defined using our *local* and *global maximum* methods. These figures complement Figure 7 of the Main text.

Figures C5 and C6 show the cumulative distribution function of the amounts given to player B, averaged per subject, separately for subjects in the Power+ preference class and those who have social preferences. Subjects with social preferences belong to the following preference classes: Social, Social&Power+, and Social&Power-. These figures complement Figure 8 of the Main text. In Table C1 we compare our subjects to those of

Charness and Rabin (2002) and Chen and Li (2009) by presenting the proportion of subjects who choose the first option in decision problems CR1-CR6. This table complements Table 5 of the Main text.

Finally, in Tables C2 and C3 we compare subjects' behavior in the Charness and Rabin (2002) task across subjects with different preferences: Standard, Power+, and social preferences (we group together Social, Social&Power+, and Social&Power- preference classes). The preference classes are defined using our *local* and *global maximum* methods. These tables complement Table 6 of the Main text. The last column reports p-values for the tests of proportions, where the alternative hypothesis is that the fraction of subjects in the Power+ preference class choosing the first option is higher than the fraction with social preferences (the null hypothesis is equality of those proportions between the two preference categories). The latter fraction is always statistically greater than the former one in all decision problems in which the second option is more efficient that the first one, i.e., in decision problems CR1-CR2, CR4-CR5, and PT1-PT2. The penultimate column reports the p-values for the tests of proportions for the Standard and Power+ preferences classes. In almost all decision problems, we cannot reject the null hypothesis that the subjects in the Power+ preference class behave similarly to those in the Standard preference class.

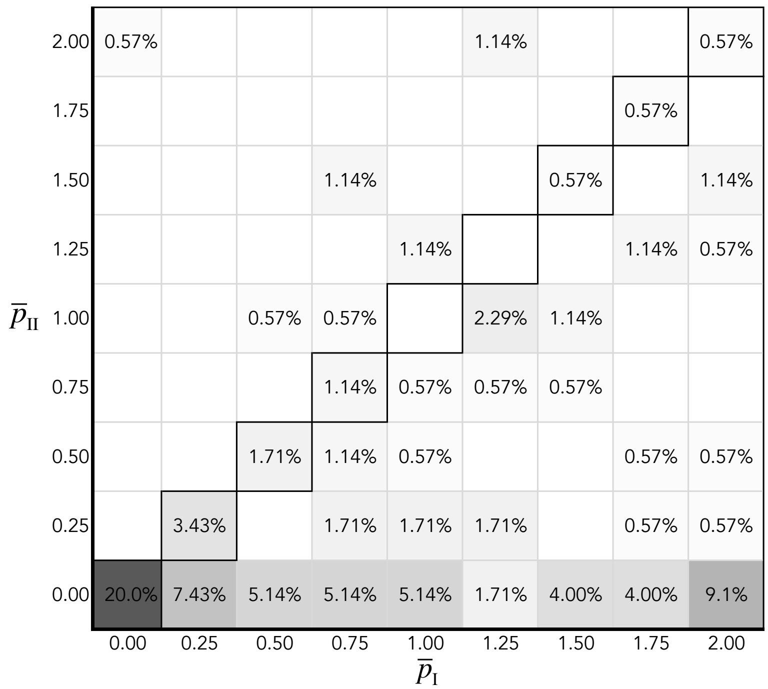

FIGURE C1. Joint distribution of the willingness to pay in Part I and Part II, $\bar{p}_{\text{I}}$ and $\bar{p}_{\text{II}}$, defined using the *Local Maximum* Method.

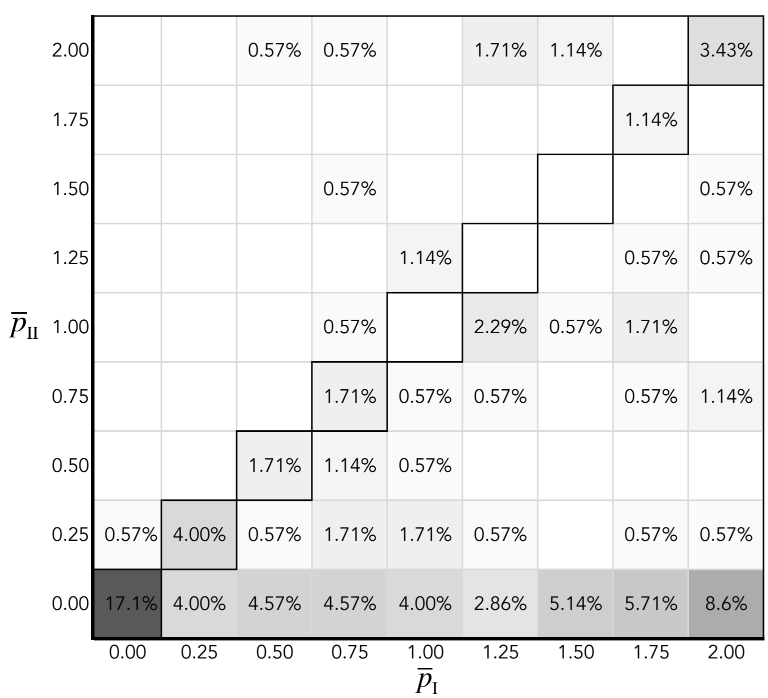

FIGURE C2. Joint distribution of the willingness to pay in Part I and Part II, $\bar{p}_{\text{I}}$ and $\bar{p}_{\text{II}}$, defined using the *global maximum* method.

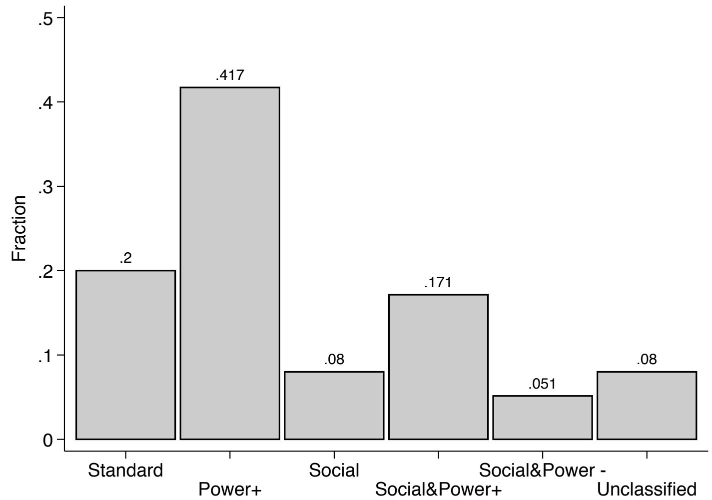

FIGURE C3. Distribution of preference classes based on the willingnesses to Pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *local maximum* method.

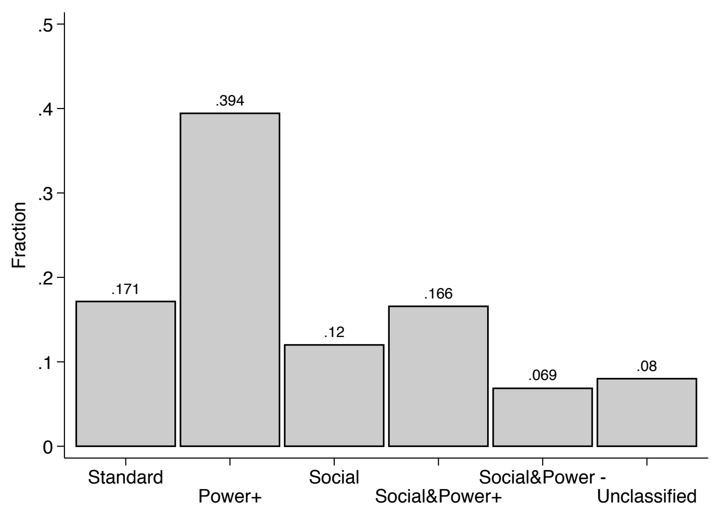

FIGURE C4. Distribution of preference classes based on the willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *global maximum* method.

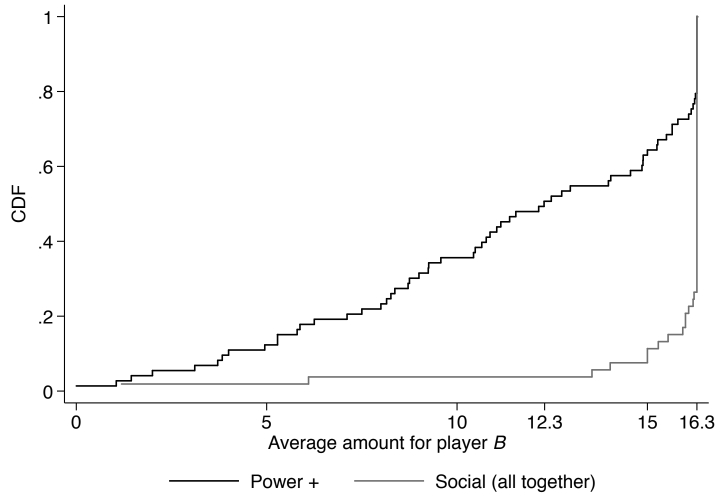

FIGURE C5. Distribution of the average amount given to player *B*, where preference classes are defined using the *local maximum* method.

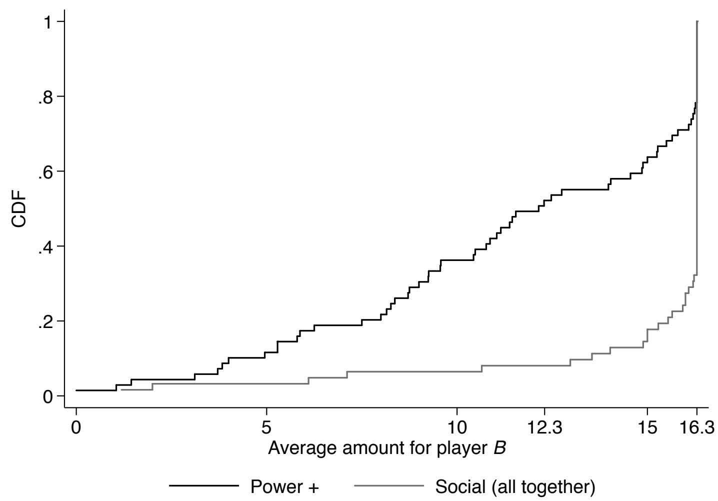

FIGURE C6. Distribution of the average amount given to player *B*, where preference classes are defined using the *global maximum* method.

TABLE C1. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task across three samples: Charness and Rabin (2002), Chen and Li (2009), and our sample.

Decision	First Option	Second Option	CR2002	CL2009	Our Subjects[a]
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$	31%	33%	30%
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$	51%	82%	53%
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$	100%	NA	97%
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$	67%	76%	81%
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$	27%	50%	59%
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$	78%	64%	78%

[a]The sample includes 175 subjects who make one or no skips across both parts of the Power Game.

TABLE C2. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task by preference class, defined using the *local maximum* method.

	Option		Preference Class			p-value for	p-value for
Decision	First	Second	Standard	Power +	Social (All)	PP = StP[a]	PP > SP[b]
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$	20%	48%	6%	0.005***	0.000***
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$	63%	74%	11%	0.236	0.000***
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$	97%	97%	98%	0.972	0.911
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$	97%	84%	68%	0.042**	0.020**
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$	80%	67%	30%	0.166	0.000***
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$	89%	77%	72%	0.145	0.262
PT1	$(10.10, 5.20)$	$(9.10, 9.10)$	86%	68%	38%	0.056*	0.000***
PT2	$(12.30, 5.10)$	$(10.10, 12.30)$	91%	75%	51%	0.048**	0.002***
PT3	$(12.30, 9.60)$	$(9.60, 12.30)$	100%	97%	100%	0.323	0.888
PT4	$(12.30, 7.80)$	$(7.80, 5.40)$	100%	97%	100%	0.323	0.888
PT5	$(6.15, 6.15)$	$(0.00, 0.00)$	100%	100%	100%	1.000	1.000

[a]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is the same for the Power+ preference class and for the Standard preference class, H0: PP = StP, H1: PP $\neq$ StP.
[b]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is greater for the Power+ preference class than for subjects with social preferences, H0: PP = StP, H1: PP > StP. *** indicates significance at 1% level, ** at 5% level, * at 10% level.

Table C3. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task by preference class, defined using the *global maximum* method.

Decision	Option First	Option Second	Preference Class Standard	Preference Class Power +	Preference Class Social (All)	p-value for PP = StP[a]	p-value for PP > SP[b]
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$	13%	49%	11%	0.001^{***}	0.000^{***}
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$	63%	72%	21%	0.364	0.000^{***}
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$	97%	97%	98%	0.908	0.688
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$	97%	86%	69%	0.104	0.013^{**}
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$	80%	71%	32%	0.351	0.000^{***}
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$	90%	77%	73%	0.126	0.289
PT1	$(10.10, 5.20)$	$(9.10, 9.10)$	87%	71%	40%	0.095^{*}	0.000^{***}
PT2	$(12.30, 5.10)$	$(10.10, 12.30)$	90%	78%	53%	0.164	0.001^{***}
PT3	$(12.30, 9.60)$	$(9.60, 12.30)$	100%	99%	98%	0.508	0.470
PT4	$(12.30, 7.80)$	$(7.80, 5.40)$	100%	98%	100%	0.346	0.912
PT5	$(6.15, 6.15)$	$(0.00, 0.00)$	100%	100%	100%	1.000	1.000

[a]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is the same for the Power+ preference class and for the Standard preference class, H0: PP $=$ StP, H1: PP $\neq$ StP.
[b]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is greater for the Power+ preference class than for subjects with social preferences, H0: PP $=$ StP, H1: PP $>$ StP. *** indicates significance at 1% level, ** at 5% level, * at 10% level.

Appendix D.

In this Appendix we show that our results are robust to using the whole sample. Even allowing for any number of skips and using different mechanisms to determine willingness to pay, we obtain the same results.

We redo all Figures and Tables that appear in Section 4.2 and Section 4.3 of the Main text, but using all 258 subjects of our sample. 126 subjects made no skips, 48 subjects skip one price across both Parts, 22 subjects make two skips, 15 subjects make three skips, and 46 subjects make four skips or more across both Parts. For those subjects who make no skips we calculate their willingness to pay in the same way as in the Main text. In Parts I and II, we use the maximum price at which they decided to pay before switching to not paying.

As in Appendix C, for those subjects who make skips, we consider two different ways to calculate their willingnesses to pay: the maximum price at which they pay before making their first skip, or 0 if they don't pay at a price of zero (local maximum method) and the global maximum price at which they pay or 0 if they don't pay at a price of zero (global maximum method). Which preference class a subject fits into depends on which method is used. For example, suppose that in Part I a subject pays all prices until \$1.00 and then pays once more at \$1.75, i.e., makes two skips, at \$1.25 and \$1.50. Suppose further that in Part II she pays only at a price of \$1.75, i.e., makes seven skips, one at each price up to \$1.50. This subject thus makes a total of 9 skips. According to our local maximum method, this subject's willingnesses to pay in Part I and Part II are \$1.00 and \$0, respectively, and she has Power+ preferences, since $p_I > 0$ and $p_{\mathrm{II}} = 0$. According to our global maximum method however, this subject's willingnesses to pay in Part I and Part II are the same and equal \$1.75, and she Social preferences, since $p_I = p_{\mathrm{II}}$.

Figures D1 and D2 show the joint distribution of the willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using our *local* and *global maximum* methods. These figures complement Figure 6 from the Main text. Figures D3 and D4 show the proportion of subjects with various preference classes based on their willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, where the preferences classes are defined using our *local* and *global maximum* methods. These figures complement Figure 7 of the Main text.

Figures D5 and D6 shows the cumulative distribution function of the amounts given to player B, averaged per subject, separately for subjects in the Power+ preference class and those who have social preferences. Subjects with social preferences belong to the following preference classes: Social, Social&Power+, and Social&Power-. These figures complement Figure 8 of the Main text. In Table D1 we compare our subjects to those of

Charness and Rabin (2002) and Chen and Li (2009) by presenting the proportion of subjects who choose the first option in decision problems CR1-CR6. This table complements Table 5 of the Main text.

Finally, in Tables D2 and D3 we compare subjects' behavior in the Charness and Rabin (2002) task across subjects with different preferences: Standard, Power+, and social preferences (we group together Social, Social&Power+, and Social&Power- preference classes). The preference classes are defined using our *local* and *global maximum* methods. These tables complement Table 6 of the Main text. The last column reports p-values for the tests of proportions, where the alternative hypothesis is that the fraction of subjects in the Power+ preference class choosing the first option is higher than the fraction with social preferences (the null hypothesis is equality of those proportions between the two preference categories). The latter fraction is always statistically greater than the former one in all decision problems in which the second option is more efficient that the first one, i.e., in decision problems CR1-CR2, CR4-CR5, and PT1-PT2. The penultimate column reports p-values for the tests of proportions for the Standard and Power+ preferences classes. In almost all decision problems, we cannot reject the null hypothesis that the subjects in the Power+ preference class behave similar to those in the Standard preference class.

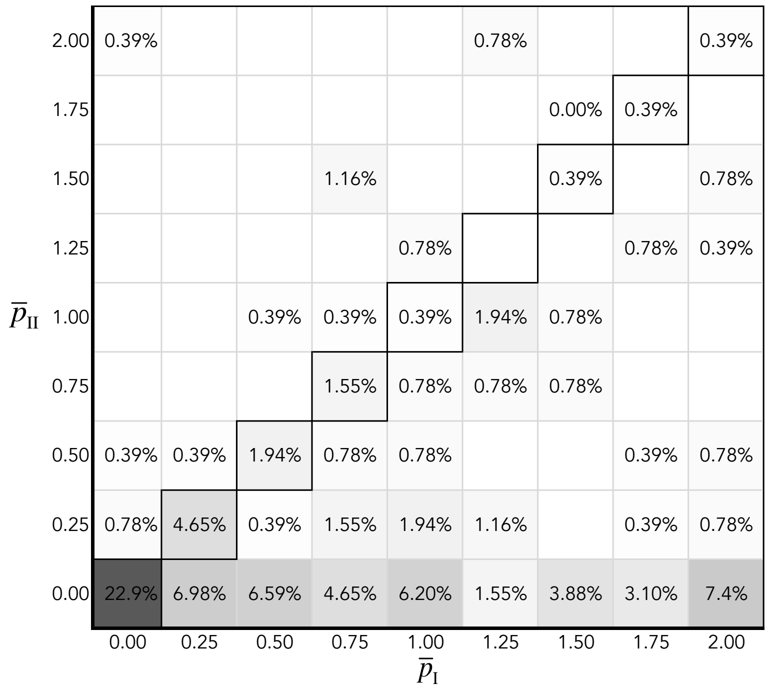

FIGURE D1. Joint distribution of the willingnesses to Pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *local maximum* method.

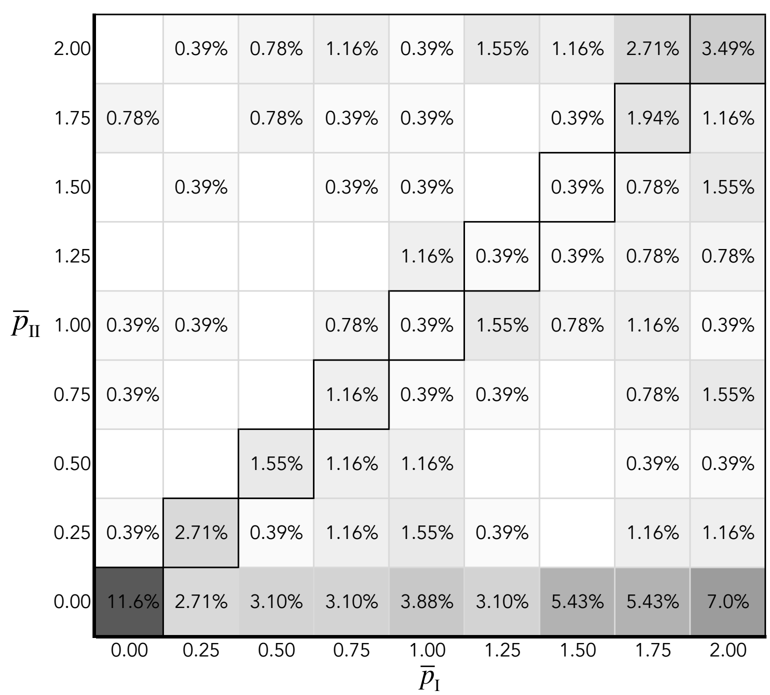

FIGURE D2. Joint distribution of the willingnesses to Pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *global maximum* method.

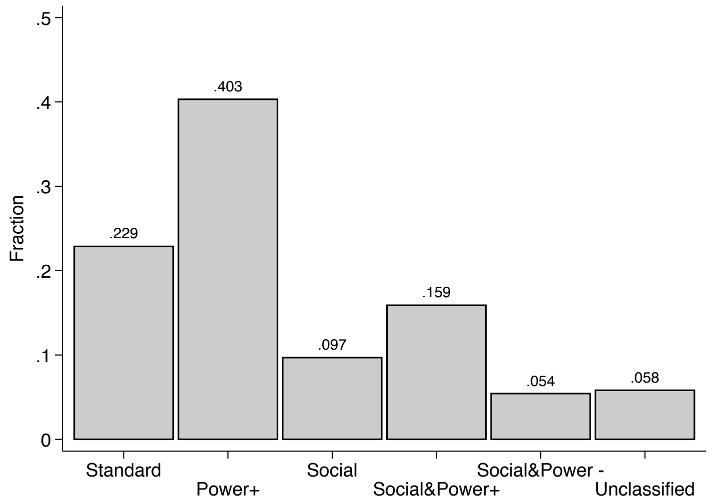

Figure D3. Distribution of preference classes based on the willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *local maximum* method.

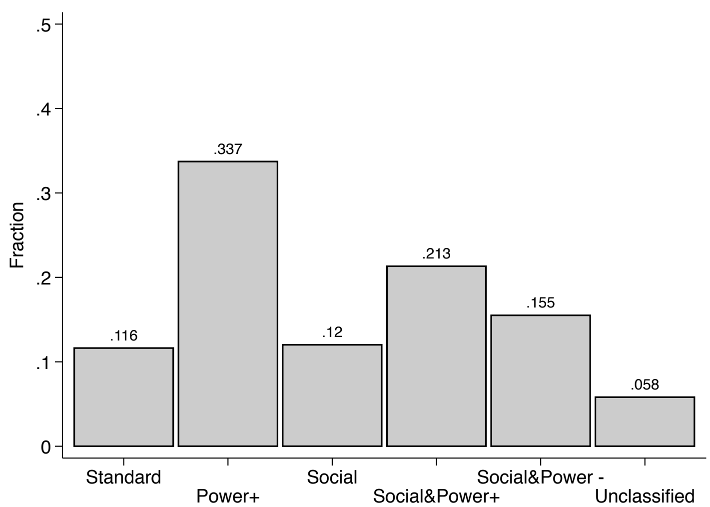

Figure D4. Distribution of preference classes based on the willingnesses to pay in Part I and Part II, $\bar{p}_{\mathrm{I}}$ and $\bar{p}_{\mathrm{II}}$, defined using the *global maximum* method.

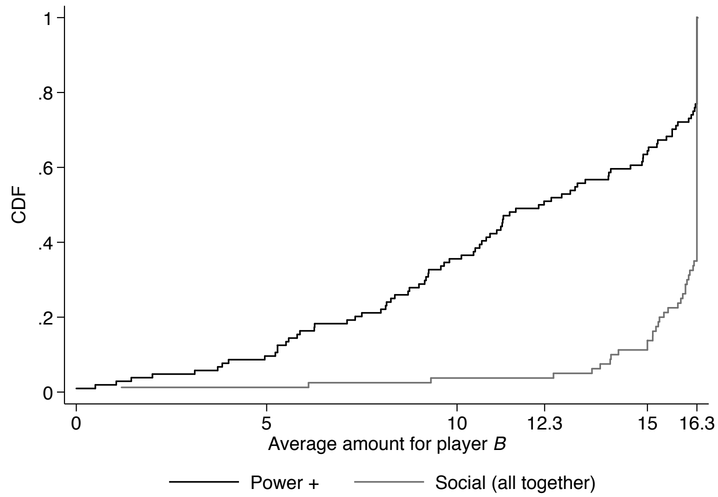

FIGURE D5. Distribution of the average amount given to player *B*, where preference classes are defined using the *local maximum* method.

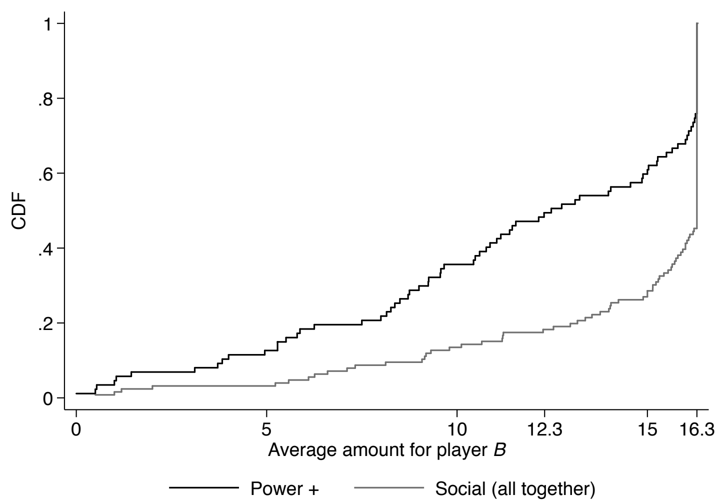

FIGURE D6. Distribution of the average amount given to player *B*, where preference classes are defined using the *global maximum* method.

Table D1. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task across three samples: Charness and Rabin (2002), Chen and Li (2009), and our sample.

Decision	First Option	Second Option	CR2002	CL2009	Our Subjects[a]
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$	31%	33%	26%
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$	51%	82%	48%
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$	100%	NA	96%
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$	67%	76%	78%
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$	27%	50%	50%
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$	78%	64%	75%

[a]The sample includes 258 subjects who make any number of skips across both parts of the Power Game.

Table D2. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task by preference class, defined using the *local maximum* method.

Decision	Option First	Option Second	Preference Class Standard	Preference Class Power +	Preference Class Social (All)	p-value for PP = StP[a]	p-value for PP > SP[b]
CR1	$(6.60, 6.60)$	$(6.60, 12.30)$	25%	38%	5%	0.091^{*}	0.000^{***}
CR2	$(6.60, 6.60)$	$(6.20, 12.30)$	63%	66%	11%	0.640	0.000^{***}
CR3	$(3.10, 12.30)$	$(0.00, 0.00)$	93%	95%	99%	0.596	0.911
CR4	$(10.50, 5.30)$	$(8.80, 12.30)$	88%	83%	66%	0.354	0.005^{***}
CR5	$(12.30, 3.50)$	$(10.50, 10.50)$	64%	57%	28%	0.337	0.000^{***}
CR6	$(12.30, 0.00)$	$(6.15, 6.15)$	76%	75%	74%	0.856	0.424
PT1	$(10.10, 5.20)$	$(9.10, 9.10)$	71%	60%	35%	0.140	0.001^{***}
PT2	$(12.30, 5.10)$	$(10.10, 12.30)$	81%	69%	48%	0.091^{*}	0.001^{***}
PT3	$(12.30, 9.60)$	$(9.60, 12.30)$	98%	98%	100%	0.917	0.894
PT4	$(12.30, 7.80)$	$(7.80, 5.40)$	100%	97%	100%	0.188	0.937
PT5	$(6.15, 6.15)$	$(0.00, 0.00)$	100%	100%	100%	1.000	1.000

[a]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is the same for the Power+ preference class and for the Standard preference class, H0: PP = StP, H1: PP $\neq$ StP.
[b]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is greater for the Power+ preference class than for subjects with social preferences, H0: PP = StP, H1: PP > StP. *** indicates significance at 1% level, * at 10% level.

Table D3. Fraction of subjects choosing the first option in the Charness and Rabin (2002) task by preference class, defined using the *global maximum* method.

	Option		Preference Class			p-value for	p-value for
Decision	First	Second	Standard	Power +	Social (All)	PP = StP[a]	PP > SP[b]
CR1	(6.60, 6.60)	(6.60, 12.30)	13%	46%	12%	0.002***	0.000***
CR2	(6.60, 6.60)	(6.20, 12.30)	63%	74%	25%	0.287	0.000***
CR3	(3.10, 12.30)	(0.00, 0.00)	97%	97%	95%	0.976	0.320
CR4	(10.50, 5.30)	(8.80, 12.30)	97%	86%	69%	0.116	0.004***
CR5	(12.30, 3.50)	(10.50, 10.50)	80%	69%	28%	0.247	0.000***
CR6	(12.30, 0.00)	(6.15, 6.15)	90%	75%	71%	0.098*	0.200
PT1	(10.10, 5.20)	(9.10, 9.10)	87%	70%	36%	0.073*	0.000***
PT2	(12.30, 5.10)	(10.10, 12.30)	90%	79%	49%	0.188	0.000***
PT3	(12.30, 9.60)	(9.60, 12.30)	100%	99%	98%	0.555	0.395
PT4	(12.30, 7.80)	(7.80, 5.40)	100%	98%	99%	0.402	0.820
PT5	(6.15, 6.15)	(0.00, 0.00)	100%	100%	100%	1.000	1.000

[a]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is the same for the power preference class and for the standard preference class, H0: PP = StP, H1: PP $\neq$ StP.
[b]The p-values are reported for the tests of proportions showing whether the fraction of subjects choosing the first option is greater for the power preference class than for the social preference class, H0: PP > StP, H1: PP > StP. *** indicates significance at 1% level, * at 10% level.