The Allures of Power

Moira Gidseg*

September 16, 2025

Abstract

In this paper, I investigate the pursuit of "power" – the ability to determine the outcomes of others – along three dimensions: who desires power, when they desire power, and what they do with power. Using a lab experiment, I classify people into four mutually exclusive "power types" based on their behavior, and I find that these power types differ in how much they spend on power, which features of power they value, and the way they respond to changes in the price of controlling their partners' outcome. Crucially, I show that many subjects value power intrinsically.

*Department of Economics, University of Arizona; mgidseg@arizona.edu.
I thank Charles Noussair, Martin Dufwenberg, and Inga Deimen. I also thank seminar participants at the Economic Science Association Conferences.

1 Introduction

"Power is not a means; it is an end."

—Orwell, *1984*

Power is a famous motivator, but *what* about power is motivating? When "power" refers to control over one's own outcomes, then there is a clear self-interest component. With power, one can directly select options which lead to higher personal utility or reduce the risk of someone else lowering one's utility. In addition, past research has found a preference for decision rights, even when it is not conducive to higher payoffs, either due to the "illusion of control" (Sloof and von Siemens, 2017) or a preference for autonomy (Bartling, Fehr, and Herz, 2014; Fehr, Herz, and Wilkening, 2013; Freundt et al., 2023; Meemann, 2023; Owens et al., 2014). However, power can also refer to control over the outcomes of *other* people. Positions granting such power are plentiful in the real-world (e.g., politicians, bureaucrats, teachers, managers, parents, and online forum moderators), yet the economics literature has largely ignored the allures of such power.

One potential allure of power is what I shall call "distributional preferences." There is much evidence in the experimental economics literature that people care about others' payoffs (e.g., Andreoni and Miller (2002), Bolton and Ockenfels (2000), Charness and Rabin (2002), and Fehr and Schmidt (1999)). I use "distributional preferences" to refer to any preferences regarding the *monetary outcomes* of oneself and others, and I do not assume a specific model of such preferences. For example, holding one's own payoff constant, a person may be altruistic (i.e., prefer to maximize others' payoffs), spiteful (i.e., prefer to minimize others' payoffs), or inequality averse (i.e., prefer others' payoffs to equal their own). Note that distributional preferences are entirely concerned with the outcome; it is irrelevant whether you, in your position as the power-holder, make the outcome occur, or if it is entirely out of your control.

In contrast to distributional preferences, "intrinsic power preferences" comprises the appeals of power which can only be obtained by being the one in power. I consider three intrinsically valuable aspects of power: "discretion," "self-recognition," and "other-recognition." Discretion is the size of the choice set from which one can select others' outcomes. Discretion may have positive or negative utility, depending on the individual: some people may relish having others at their mercy, while other people may be burdened by a sense of responsibility. "Self-recognition" is the extent to which the power-holder herself knows her power. Someone who knows that she is the pivotal voter may enjoy her power more than someone who is uncertain whether she is pivotal. "Other-recognition" is the extent to which others know about a power-holder's power. A Congress member who is recognized by his constituents may enjoy his power more than a no-name bureaucrat.

I investigate the allures of power with policy in mind. People are not randomly sorted into positions of power, and I suspect that they do not use power randomly, either. Instead, people will differ in their

distributional preferences and attitudes toward discretion and recognition, and they will self-select (such as run for office, interview for a job position, volunteer, etc.) into positions that give them the sort of power they desire. It is scientifically interesting to measure the degree to which people value these various allures. Further, there may be opportunities for intervention if people's attitudes towards these allures are correlated with their behavior as power-holders. If, for example, benevolent people are more inclined than malevolent people towards positions which are high in recognition by others, then we should design important positions to interact heavily with the public. Demographics may also play a role here: if, for example, women tend to be more benevolent, then we may attract better power-holders by advertising positions using female actors or language that appeals to women.

In this paper, I use an experiment to address questions that can be grouped into three broad categories: "when do people desire power," "what do they do with it," and "who desires it." I first classify subjects into four mutually-exclusive "power types" based on how they behave in a position of power. In terms of when people desire power, I separate the appeals of distribution, discretion, and recognition, as well as consider the importance of power's price and direction (i.e., whether one can increase or decrease outcomes). In terms of what people do with power, I measure the prevalence of each power type, when decisions are motivated by intrinsic power preferences, and how often subjects buy power which they do not "use." Finally, in terms of who desires power, I examine how power types differ in the amount they spend on power, as well as how well personal characteristics predict power type and spending.

To measure how much people value power, I conduct an experiment in which subjects play several rounds of the "Monetary Power Game," inspired by Pikulina and Tergiman (2020), as well as a novel control, the "Points Power Game." In both games, subjects are paired up, where one is randomly assigned to the role of the potential power-holder and the other is a passive player. In Pikulina and Tergiman's Power Game, potential power-holders have the choice to buy, for a fixed fee, the power to pick the experimental earnings of another player from the interval \$0 to \$16.30. Subjects played the Power Game multiple times, varying the fixed fee between rounds and thereby allowing the authors to estimate how much subjects are willing to pay for the interval \$0 to \$16.30. All subjects were asked to make decisions, but the true power-holder (i.e., the one whose decision determines the payoff of herself and the player she was matched with) was chosen randomly for each matched pair of players, without the players ever finding out whose decision counted.

In designing this game, Pikulina and Tergiman (2020) removed self-interest as a motivation for obtaining power – something which so often acts as a confounding factor in real-world scenarios. Their study established not only that there is a portion of the population which values having power over others, but also that some of them value power intrinsically. In addition, they briefly tested whether the size of the power-holder's choice set alters the willingness to pay for power, and they found that it does. However, they only do one

comparison, in which they simultaneously shrink the choice set and remove the option to increase the other person's earnings. In this paper, I further investigate the importance of the size of the choice set and the presence of options to increase or decrease the other's earnings. I do this by introducing two key changes to their design in the Monetary Power Game.

First, subjects no longer have a fixed fee for a fixed interval of power; instead, they can choose how much to spend on power, and the menu from which they can pick their partner's experimental earnings expands proportionally. Subjects still play the game multiple times, but what varies is how much the menu expands for each dollar the subject pays. This includes both the magnitude of expansion and the direction, so that in some cases, subjects can only increase their partners' earnings compared to the no-power case, while in other cases, subjects can only decrease those earnings, and in still other cases subjects can do both. By letting subjects choose how much power to buy, unlike Pikulina and Tergiman (2020), subjects also control whether they buy more power than they "use": they can either buy exactly as much power as is needed to choose their desired earnings for their partners, or they can buy more power than necessary. The latter case would be strong evidence that discretion has intrinsic value, rather than solely value from achieving desired earnings for one's partner.

Second, I vary between subjects how much the passive player and the power-holder herself know about the power-holder's position. The control treatment mimics Pikulina and Tergiman (2020), where each subject makes decisions as the power-holder, knowing that there is an equal chance that she or her partner may be the true power-holder. In the "high information" and "low information" treatments, the power-holder knows that she is the power-holder with certainty. The two information treatments differ in what the passive player learns: in the "low information" treatment, the passive player only knows that another player influenced his earnings, whereas in the "high information" treatment, the passive player will be shown the instructions given to the power-holder and the choice the power-holder made in the round which determined the passive player's earnings.

In order to separate distributional preferences from intrinsic power preferences, I have subjects participate in two additional games. The first is the "Points Power Game." For the remainder of the paper, I will shorten "Monetary Power Game" and "Points Power Game" to just "Monetary Game" and "Points Game," respectively. As the name suggests, the Points Game is identical to the Monetary Game except that the power-holder is only able to change her partner's "points" rather than experimental earnings. These points have no impact on earnings, so distributional preferences are irrelevant to this decision. The second method of separating distributional preferences from intrinsic power preferences is the "Binary Choice Game," inspired by Pikulina and Tergiman (2020). Here, the power-holder chooses between two options, each specifying the earnings for herself and her partner. Since the power-holder's menu is fixed at two options, intrinsic power

preferences are irrelevant to this decision, leaving only distributional preferences. I compare each subject's behavior in the Monetary Game to her behavior in these controls in order to identify which decisions are driven by intrinsic preferences for power.

The personal characteristics I gather from subjects are their risk aversion, Big Five personality traits (extraversion, agreeableness, conscientiousness, neuroticism, and openness), preference for authoritarianism, prior experience in economics experiments, and demographics. As mentioned earlier, the goal here is to determine whether personal characteristics can predict power type, which would allow tailoring advertisements for positions of power (e.g., managers, teachers) to the people most likely to be benevolent power-holders.

I find several pieces of evidence that subjects value power intrinsically. In both the Monetary Game and the Points Game, subjects often bought more power than they use, implying that unused options have intrinsic value. In the Binary Choice Game, subjects indicated that they preferred the outcome they could have achieved by not buying power rather than the outcome they bought in the Monetary Game, suggesting once again that the outcome was not the key motivator. Finally, many subjects bought as much power in the Points Game as in the Monetary Game, implying that control over worthless points is as valuable as control over monetary outcomes. Because neither self-recognition nor other-recognition affected behavior, I attribute this intrinsic value of power to liking discretion. In contrast to the abundant evidence of positive intrinsic power preferences, subjects who dislike power seem to be very rare.

In terms of power types, the best type to be matched with as a passive player was benevolent: someone who, whenever she bought power, never decreased her partner's earnings. These subjects comprised 28% of the sample. As one might guess, these subjects chose the highest average earnings for their partners compared to other types. However, their behavior seems to be motivated by discretion rather than distribution. These subjects like discretion, especially discretion to increase their partners' outcomes, but they are willing to buy just as much power in the Points Game as in the Monetary Game.

The most common type was capricious: someone who chose to increase and decrease her partner's earnings at least once each. These subjects comprised 37% of the sample, and made larger magnitude changes to their partners' earnings than benevolent types did. Despite sometimes decreasing their partners' earnings, these types are overall slightly beneficial to their partners. As one might expect from their inconsistent choices, this type also buys the most unused discretion, so they are the most prone to buying power for discretion's sake alone.

The least common type was malevolent: someone who, whenever she bought power, never increased her partner's earnings. These subjects comprised 10% of the sample, but they spent as much on power as capricious types and had as large of an impact on their partners' earnings. They like discretion, but they also care about the distributional consequences: those who played the Monetary Game first bought much

less power in the Points Game.

Lastly, selfish types are those who bought power no more than once. These subjects comprised 24% of the sample. If we look only at rounds when subjects bought power, the 35% of selfish subjects who bought power once made the largest difference to their partners' earnings, and they always chose to increase. In the Binary Choice Game, they also were the least likely to choose an outcome which was worse for both players. While selfish types are unlikely to pursue a position of power of their own volition, evidence suggests that they may be the best choice to have as power-holder – provided we have a way of selecting power-holders based on type. Unfortunately, personal characteristics are not helpful here: the only robust effect on behavior was that subjects who scored higher on openness or agreeableness tended to spend more on power.

This study shares connections with many others in the social preferences literature. For example, García-Gallego et al. (2019) designed and conducted the Heaven's Dictator Game, in which the heaven's dictator gets to decide the experimental earnings of another subject. Unlike the standard dictator game, the heaven's dictator's earnings are fixed, so there is no trade-off between the heaven's dictator's own earnings and that of the other person. Their results provide some insight on "What do power-holders do with power?" but cannot speak to "who desires power" or "when do they desire it."

There is some overlap between power and paternalism, in that paternalistic feelings may drive a person to seek power. Existing literature already shows that people are willing to engage in paternalism, i.e., restricting the freedom of others to make their own decisions, out of concern for the welfare of those they are restricting (Ambuehl et al., 2021; Bartling, Cappelen, et al., 2023; Grossmann, 2024, 2025). Note that in the Monetary Game and Points Game, paternalism is not applicable, because the potential power-holder's partner has no choice *regardless* of whether the potential power-holder buys power or not. Of course, this is an artificially created environment, and in many real-world situations, power-holders may choose whether to dictate outcomes or leave them up to individuals. In such cases, paternalism can be one motivator for restricting outcomes, but other motivations are possible. For example, power-holders may seek to improve others' outcomes not out of concern for others' welfare *per se*, but rather because they enjoy exerting power. Power-holders may also have spiteful preferences which motivate them to choose outcomes *worse* than the ones they believe individuals would choose for themselves. If potential power-holders are motivated by paternalism, heterogeneity of beliefs can introduce heterogeneity in behavior: when obtaining power is costly, a paternalistic person may only choose to buy power if she believes that others will not make "good" choices without her intervention. The Monetary Game and Points Game are designed to eliminate such heterogeneity of beliefs.

Dana et al. (2006) investigated the importance of the other player's knowledge of the power-holder's decision. In their experiment, subjects are asked to split \$10 in a dictator game. However, after making

their decision and before the money is actually sent to the other player, the dictator is given the option to opt out of the dictator game, in which case the dictator gets $9 while the other player gets nothing. In one treatment, if the dictator chose to go through with the dictator game, both the money and the instructions for the game would be transferred to the second player. In a second treatment, if the dictator went through with the dictator game, the second player would not be given the instructions. The frequency of opting out was much lower in the second treatment than in the first, indicating that subjects derived negative utility from the other player learning of their position/decision as dictator. On the other hand, fame is often thought to be one of the perks of being a politician or similar power-holder. The information treatments in my experiment contribute to the literature on the way active players' behavior changes depending on the information given to the passive players.

There is an existing literature on the value of autonomy (Bartling, Fehr, and Herz, 2014; Fehr, Herz, and Wilkening, 2013; Freundt et al., 2023; Owens et al., 2014). It is well-established that having control of one's own outcome is a good. However, enjoying control over one's own outcome need not imply that one enjoys control over other people's outcomes. In many real-world contexts, the two go hand-in-hand: there are many positions which grant power over others' outcomes, and if one does not occupy that position oneself, then another person will. Since autonomy is not a focus of this paper, I remove any autonomy considerations from the decision to seek power in my experiment: the potential power-holder is certain of what her own outcome will be whether or not she buys power.

This paper also contributes to the literature on the relationship between personality and behavior in social contexts. Many articles have examined the relationship between the Big Five Inventory dimensions and altruism, but results are mixed. Some papers find that agreeableness is positively associated with altruism (Baumert et al., 2014; Becker et al., 2012; Ben-Ner, Kramer, and Levy, 2008; Koole et al., 2001; LePine and Van Dyne, 2001) and others find no (or even negative) effects (Ben-Ner and Kramer, 2011; Visser and Roelofs, 2011). The Monetary Game provides the power-holder with an opportunity to lower her own earnings in exchange for increasing her partner's earnings, and I investigate whether doing so is correlated with the Big Five Inventory traits, among other subject characteristics.

The rest of the paper is organized as follows. Section 2 describes the experimental design. Section 3 describes the conceptual framework. Section 4 explains how I will detect intrinsic preferences and then lays out the hypotheses and exploratory analyses. Section 5 describes the data. Section 6 reports and discusses the results. Section 7 concludes.

2 Experimental Design and Protocol

2.1 The Games

In the Monetary Power Game, there is one active player ("A") and one passive player ("B"). Both players start with \$12.30. Player A chooses how much of her payoff to spend, $s \in [0, 2]$ dollars, in exchange for the ability to choose the earnings for player B. If A chooses $s = 0$, the game ends, and both A and B get earnings of \$12.30. If instead A chooses $s > 0$, then she can change the earnings for player B by $y \in [-ds, us]$, where $d \geq 0$ is a decrease factor and $u \geq 0$ is an increase factor (u for "up"), both revealed to A before choosing s. The final earnings are $\$12.30 - s$ for A and $\$12.30 + y$ for B.

The Points Power Game[1] is very similar: A is active, B is passive, and both start with \$12.30. Player A chooses how much to spend, $s \in [0, 2]$, except now it is in exchange for the ability to choose player B's *points*. Both A and B start off with 12.30 points. These points do not affect experimental earnings in any way; one's own points are merely displayed alongside one's earnings at the end of the game. As above, if A chooses $s = 0$, the game ends with both A and B getting \$12.30 and 12.30 points. If A instead chooses $s > 0$, then A can change the points for player B, $y \in [-ds, us]$, where $d, u > 0$ are revealed to A before choosing s. The final earnings are $\$12.30 - s$ for A and \$12.30 for B. The final points are 12.30 for A and $12.30 + y$ for B. In the experiment, for both the Monetary Game and the Points Game, s and y were discretized so that each must be in increments of 5 cents. A summary of the notation for these games is shown in Table 1.

Table 1: Notation for the Monetary Game and Points Game

Symbol	Meaning
s	amount of own earnings (in dollars) that power-holder spends on power
y	amount by which power-holder changes her partner's outcome[a]
u	increase factor, i.e., multiplier that determines how high y can be for given s
d	decrease factor, i.e., multiplier that determines how low y can be for given s

Note: This table describes the notation used for the Monetary Game and Points Game.
[a]This outcome is earnings (in dollars) in the Monetary Game and points in the Points Game.

These games have several convenient features. First, the only decision that affects player A's earnings is deciding how much to spend on power. This presents a clear trade-off between self-interest and power, allowing us to measure the value of power in terms of willingness to pay. Second, A can choose how much power to purchase. This differs from Pikulina and Tergiman (2020), where subjects always purchased the same fixed menu of options for y. This change allows us to observe whether subjects choose to buy more power than is required to select their preferred y. Since there is no new information received between choosing

[1]I thank my advisor, Charles Noussair, for the idea of the Points Game.

s and y, and the time-gap is minimal, buying more power than required would be strong evidence of a preference for discretion itself. Third, by allowing cases where $d \neq u$, we can observe whether subjects value power differently when it involves only increasing another's earnings above the no-power case, as opposed to only decreasing or having both options. Finally, since B is a passive player, A's actions should be free of strategic or reciprocity considerations.

2.2 Protocol for the Experiment

The experiment consisted of four parts and a survey. Subjects were allowed to advance at their own pace through the experiment, but no one was allowed to receive their payment and leave until everyone was finished. Each session lasted roughly one hour. Average earnings were \$12.08 plus a \$5 show-up fee for those who made decisions (i.e., all subjects in the control treatment, and subjects acting as player A in the low/high information treatments), and \$12.54 for purely passive players (i.e., subjects acting as player B in the low/high information treatments). All data collection was done on computers in the University of Arizona's Economic Science Laboratory using z-Tree software (Fischbacher, 2007). Subjects were paid for one round from one part at random, so that each round had an equal probability of being selected. The round that counted for payment was the same for both A and B, and subjects knew this fact, though not which round it was.

For each part, after receiving the instructions and before making any decisions, subjects were given a quiz to test their understanding. Subjects needed to get every question correct to proceed. The quiz consisted of only "fill-in-the-blank" or "select all that apply" questions, severely limiting the ability for subjects to proceed by guessing. The quiz questions are available in the Appendix.

After finishing part 1, subjects were informed that they were anonymously paired at random with another subject. In parts 2-4, subjects made decisions about both their own earnings and that of their partner. In the control treatment, the subject's partner was another subject in the same session, and either the subject's or his partner's decisions would determine the earnings for both of them, with equal probability. Both subjects were asked to make decisions as if they were the one who determined earnings; subjects were never told whether their decisions were the ones that counted. This was the structure in Pikulina and Tergiman (2020), which has the potentially important property that one has the same probability of having power as that of the person that one may have power over. To see whether lacking the information that one is the definitive power-holder dilutes the value of power (and thus willingness to pay), I also conducted two treatments where the subject knows she is the only one with power between her and her partner. I next describe these treatments.

In the "high information" and "low information" treatments, the power-holder's partner (B) was a subject in a later experiment. The power-holder (A)'s decision was guaranteed to determine the earnings for both herself and B. In the control and both information treatments, B was shown his points and earnings. However, in the "low information" treatment, all B was told is that a subject in another experiment influenced his earnings, and in the "high information" treatment, B was given a copy of A's instructions, shown which part was selected for payment, and shown what A did in the round that counted for payment. By varying what B knows about A's decision, and telling A about what B will know, we can see the effect of B's information on both A's willingness to pay for power and the outcome A chooses for B. The Economic Science Lab where all sessions were conducted has a long history of a strict no-deception policy, and this policy is also displayed prominently within the lab, so when subjects are told about the information structure, it is credible.

To control for and evaluate ordering effects, half of the subjects did part 2 before part 3, while the other half did part 3 before part 2. Subjects knew there were four parts from the beginning of the experiment, but did not learn the details until instructions were distributed at the beginning of the corresponding part. Instructions can be found in the Appendix.

2.2.1 Part 1 (Risk Aversion Measure)

Part 1 consisted of 5 rounds in which subjects were asked to choose between guaranteed earnings and a gamble which yielded either \$4.30 or \$20.30 with equal probability. The guaranteed earnings could be \$4.30, \$8.30, \$10.30, \$12.30, or \$14.30. The order of the guaranteed earnings was random, but all subjects were asked to make decisions for each guaranteed earnings option. Whether the guaranteed value was shown as the left or right option was also randomized. The lowest guaranteed earnings for which subjects chose the guarantee instead of the gamble was used as a measure of risk aversion. This measure was then converted to a scale of 0-5, where 0=least risk averse and 5=most risk averse.

Note that there is no risk involved in the Monetary and Points Games. I measure risk aversion only because higher risk aversion may be positively correlated with desire for power. The intuition is that someone may be risk averse because she dislikes uncertainty over her payoff; and may instead prefer control over her own payoff. Further, a desire for control over one's own payoff may also be associated with a desire to control one's environment in general – including the outcomes of others.

2.2.2 Part 2 (Monetary Game)

Part 2 consisted of 11 rounds in which subjects played the Monetary Game. The rounds varied solely in the values of the increase factor u and decrease factor d. All combinations used in the experiment are shown in Table 2. The order of the rounds was random, but all subjects were asked to make decisions for each

combination.

Table 2: Parts 2-3 Increase and Decrease Factors

Round[a]	Increase factor u	Decrease factor d
1	0.25	0.25
2	0.5	0.5
3	1	1
4	2	2
5	4	4
6	0.5	0
7	1	0
8	2	0
9	0	0.5
10	0	1
11	0	2

Note: For spending s, player A could choose player B's outcome from the range $[12.30 - ds, 12.30 + us]$.
[a]The order of rounds was randomized for each subject.

2.2.3 Part 3 (Points Game)

Part 3 consisted of 11 rounds in which subjects played the Points Game. The rounds varied solely in the values of the increase factor u and the decrease factor d. The combinations used in the experiment were the same as in Part 2. The order of the rounds was randomized separately from the order in Part 2; all subjects were asked to make decisions for each combination.

2.2.4 Part 4 (Binary Choice Game)

Part 4 consisted of 23 rounds in which subjects were asked to choose between two options specifying their own earnings and those of their partner. The choices for 12 rounds (the "common rounds") were the same for all subjects and are shown in Table 3. For the remaining 11 rounds, one of the choices was for both players to receive \$12.30, while the other option depended on the subject's choices in a round of part 2. Each of the 11 rounds of part 2 has an analogous round in part 4. For some round in part 2, let $\tilde{s}$ be the earnings the subject paid, $\tilde{y}$ be the amount by which the subject changed her partner's earnings, and $\tilde{u}$ be the increase factor. If $\tilde{s} > 0$, then the analogous part 4 round's second option offered $\$12.30 - \tilde{s}$ for the subject and $\$12.30 + \tilde{y}$ for her partner – that is, the same outcome that the subject chose in part 2. If instead $\tilde{s} = 0$, then the second option offered \$11.30 for the subject and $\$12.30 + \tilde{u}$ for her partner.

Note that in each of these 11 rounds of part 4, both options presented to the subject were attainable in the corresponding round of part 2. Also, one of the options is always what the subject actually chose

Table 3: Part 4 Common Rounds

Round[a]	Option 1	Option 2
1	(10.90, 10.90)	(10.90, 16.60)
2	(10.90, 10.90)	(10.50, 16.60)
3	(14.80, 9.60)	(13.10, 16.60)
4	(14.40, 9.50)	(13.40, 13.40)
5	(16.60, 7.80)	(14.80, 14.80)
6	(16.60, 9.40)	(14.40, 16.60)
7	(16.60, 13.90)	(13.90, 16.60)
8	(16.60, 4.30)	(10.45, 10.45)
9	(7.40, 16.60)	(4.30, 4.30)
10	(16.85, 17.10)	(16.60, 16.60)
11	(10.45, 10.45)	(4.30, 4.30)
12	(16.60, 12.10)	(12.10, 9.70)

Note: For each tuple, the first number represents the subject's own earnings, and the second number represents the subject's partner's earnings (both in dollars).
[a]The order of rounds was randomized for each subject.

as an outcome in part 2 (option 1 for those who did not pay, and option 2 for those who did pay). Thus, contrasting behavior between part 2 and part 4 allows us to determine whether the subject's choices in part 2 were due to distributional preferences or intrinsic power preferences. If a subject chooses different earnings in part 4 compared to the corresponding round of part 2, then her choice in part 2 cannot be driven solely by preferences over earnings – instead, it indicates at least some preference about power *per se*.

The order of the rounds was random, as was which option was displayed on the left, but all subjects were asked to make decisions for each. The options are based on part 2 of Pikulina and Tergiman (2020)'s experiment, with the payoffs in the common rounds increased by \$4.30. Shifting up the payoffs this way has two merits: avoiding \$0 as an option (and thus the zero payoff effect) and making these payoffs more comparable with payoffs from the rest of the experiment.

2.2.5 Survey

After completing parts 1-4, subjects were given a survey. In order to investigate the correlation between personality and preferences for power, subjects first completed a 10-question version of the Big Five Inventory from Rammstedt and John (2007), which measures personality on a 10-point scale along each of five dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness.

To investigate the correlation between authoritarianism and preferences for power, subjects then completed a measure of authoritarianism from Engelhardt et al. (2023). This measure asks subjects to pick which of two traits they think is more important for a child to have, for 8 pairs of traits. Compared to

other measures of authoritarianism, this measure has the benefit of being exogenous to political attitudes and more temporally stable for an individual.

The remainder of the survey included questions about demographics, whether anything was confusing, what the subject thought the goal of the experiment was, and what motivated the subject's decisions in parts 2-4. The exact wording of the survey questions can be found in the Appendix.

3 Conceptual Framework

3.1 Utility Components

I assume that a person's utility function is composed of five additively separable components: (1) her distributional preferences, which are a function of only the monetary payoffs to herself and her partner; (2) her positive discretion, which scales with the range of outcomes she can choose for her partner, considering only those outcomes that constitute "helping" her partner; (3) her negative discretion, which scales with the range of outcomes she can choose for her partner, considering only those outcomes that constitute "hurting" her partner; (4) her self-recognition, which is a function of the probability she is a power-holder and the range of outcomes she can choose for her partner; (5) her other-recognition, which is a function of what her partner knows about her position as a power-holder and the range of outcomes she can choose for her partner. I do not assume that all subjects care about all five components; the utility from any of these components could be zero.

$$U = U_{distribution} + U_{positive\text{-}discretion} + U_{negative\text{-}discretion} + U_{self\text{-}recognition} + U_{other\text{-}recognition}$$

I make two key assumptions about distributional preferences. The first is that they are stable across decisions. This is made more plausible by the fact that the power-holder's partner is a passive player, rendering reciprocity concerns irrelevant. The second assumption is that, holding her partner's earnings constant and conditional on having weakly lower earnings than her partner, the power-holder's distributional utility is strictly increasing in her own earnings. This still allows subjects to have a variety of preferences, including spiteful, altruistic, welfare-maximizing, selfish, inequality averse, and maximin. In fact, I am not aware of any social preference model which does not fulfill this assumption.

Discretion is evaluated in the middle of the game. If discretion were defined as the range of feasible outcomes at the beginning of the game, then discretion would be a constant property of the game, since it would not depend on anyone's actions. If it were instead defined as the range of feasible outcomes at the end of the game, then discretion would always be zero, because only one outcome will be realized at the end

of the game. Thus, in order for discretion to be relevant to players' actions, it must be defined by the range of options at some interim point in the game.

In the Monetary Game and Points Game, discretion is evaluated once the power-holder buys power and before she must decide how to use it. If the power-holder spends s on power, then she has discretion $us + ds$. I assume that the power-holder considers "helping" to be giving her partner anything above his no-power outcome, and "hurting" to be anything below his no-power outcome. This means that positive discretion is us and negative discretion is ds. In the Binary Choice Game, there is only one stage, and both options within that stage immediately result in a single outcome. Thus, the Binary Choice Game has no discretion no matter which option the power-holder chooses. The risk aversion measure in part 1 also has no discretion because that part does not affect anyone else's outcomes. Since the only games with discretion are the Monetary Game and Points Game, those are also the only games for which I need to assume how subjects will classify positive versus negative discretion. Note that since the power-holder and her partner have the same no-power outcome, nothing would change if we instead assumed that the power-holder used her own no-power outcome as the separation between positive and negative discretion.

The risk aversion measure only concerns the subject's own payoff, so the subject has no power during this part. Once subjects are paired up, their self-recognition and other-recognition depends on their treatment assignment. For a given amount of power bought, those in the high or low information treatments have higher self-recognition than those in the control treatment, because they have a 100% chance of being the power-holder instead of a 50% chance. Those in the high information or control treatments have higher other-recognition than those in the low information treatment, because their partners know all of the games that the power-holders play. I do not attempt to quantify the difference in recognition between the treatments; I only claim that recognition is higher in one treatment than in another. Utility from both self-recognition and other-recognition increases in the range of options for one's partner's outcome: being recognized as having a lot of power is more valuable than being recognized as having only a little. This means that recognition will not affect decisions in the Binary Choice Game, for the same reason that discretion will not affect those decisions.

The only difference between the Monetary Game and the Points Game is that the power-holder cannot change the earnings of her partner in the Points Game. The power-holder can still buy power, and that comes with the same discretion and recognition as in the Monetary Game. However, buying power no longer has the distributional *incentives* that were present in the Monetary Game. It still has distributional *effects*: the power-holder lowers her own monetary payoff when she buys power. However, by assuming that distributional preferences are strictly increasing in one's own payoff, I rule out the possibility that lowering one's own payoff is itself a benefit. Within my theoretical framework, therefore, a subject will only

buy power in the Points Game if she values power for non-distributional reasons – that is, liking discretion or recognition. A summary of which components of the utility function are relevant for each part of the experiment is shown in Table 4.

Table 4: Factors Present by Experiment Task

Task	Distributional	Discretion	Self Recognition	Other Recognition
Risk aversion	Yes	No	No	No
Monetary Game	Yes	Yes	Depends[a]	Depends[b]
Points Game	Partial[c]	Yes	Depends[a]	Depends[b]
Binary Choice Game	Yes	No	Depends[a]	Depends[b]

Note: This table describes the considerations which influence behavior in each part of the experiment.
[a]Self-recognition is high for subjects in the high information and low information treatments. It is low for subjects in the control treatment.
[b]Other-recognition is high for subjects in the high information and control treatments. It is low for subjects in the low information treatment.
[c]Buying power will reduce one's own payoff, but subjects cannot change their partners' payoffs.

3.2 Power Types

When analyzing subjects' behavior, I classify each person into one of four mutually exclusive "power types," based on their choices in the Monetary Game: selfish, benevolent, malevolent, or capricious. These labels are consequentialist rather than theoretical: from a policy perspective, power-holders who use power to improve the lives of others are generally more desirable than those that use power to reduce the well-being of others. It is therefore useful to examine power-holders separately depending on their use of power. However, even if several people all reliably use power to improve the lives of others, that does not mean that they all share the same underlying motives. Some may be motivated solely by distributional preferences, while others may value a mixture of discretion and other-recognition, etc. This is what I mean by power types not being theoretical concepts: they do not map to the motivations identified in my theoretical framework. Next, I will define the power types and discuss what we can infer about their distributional preferences.

"Selfish" subjects are those who bought power in one or none of the rounds of the Monetary Game. Here, "selfish" means being purely self-interested or self-focused. I label subjects who never bought power as selfish because that is what a subject who is only concerned with maximizing her own earnings would do – we do not need to appeal to any other motivations to explain her behavior. It is possible that a selfish subject cares very weakly about distribution, discretion, or recognition. It is also possible that she cares strongly, but her favorite distribution is the no-power outcome, she actively dislikes having discretion, or she actively dislikes recognition of either type, or some combination of all of the above. We can only rule out that such subjects place strong positive value on the sum of her distributional, discretion, and recognition preferences.

I decide to also classify subjects who bought power a single time as selfish, because buying power only once may be a mistake or driven by curiosity (although the practice round aims to satiate such curiosity before subjects make real decisions). The decision to allow subjects to make one mistake is arbitrary, so I also run all of my analyses with two other definitions of selfish: (1) never bought power at all or (2) bought power two times or less. Unless otherwise mentioned, my results remain qualitatively unchanged under these alternative definitions.

A "benevolent" subject is one who (1) is not selfish, (2) increased her partner's earnings at least once, and (3) never chose to decrease her partner's earnings. This is the most altruistic type of power-holder, and she can be relied upon to not abuse her power by hurting others. Based on criterion 2, we can infer that the subject does not have spiteful or inequality averse preferences. This is because no matter how much a subject spends on power, she has the choice to not change her partner's earnings. A spiteful person would prefer to leave her partner's earnings unchanged rather than increase them. This is also true for an inequality averse person, because when the no-power outcome is both players getting the same earnings, increasing her partner's earnings will always increase inequality.

A "malevolent" subject is one who (1) is not selfish, (2) decreased her partner's earnings at least once, and (3) never chose to increase her partner's earnings. We can infer that a malevolent subject does not have altruistic distributional preferences, because she always had the option to not change her partner's earnings rather than decreasing them. Unlike benevolent types, however, we cannot categorically rule out inequality aversion, because once she has bought power, the power-holder has less than her partner, and thus decreasing her partner's earnings also decreases inequality.

A "capricious" subject is one who (1) is not selfish, (2) increased her partner's earnings at least once, and (3) decreased her partner's earnings at least once. We can infer that a capricious subject does not care about her partner's payoff, because all rounds start with the same no-power outcome, and buying power always lowers her own payoff; there is no distributional reason that she would choose to raise her partner's earnings in some rounds but lower them in others. And once we rule out distributional incentives to buy power, we can conclude that these subjects place positive intrinsic value on power, although we cannot yet discern whether this value comes from discretion, self-recognition, or other-recognition.

Note that this method of classifying power types still leaves some behavior unclassified. If a subject buys power often enough to not be classified as selfish yet never changes her partner's earnings, then she cannot be considered benevolent, malevolent, or capricious. In practice, no subjects did this, but how could such a subject be classified? She must have no preferences regarding her partner's outcome, valuing power only intrinsically. The alternative is that she has distributional preferences such that she gains utility from unilaterally lowering her own earnings, but that would contradict the assumption that distributional

preferences are strictly increasing in one's own payoff. Thus, such subjects would best be grouped with capricious types, who similarly do not value their partners' earnings.

4 Research Questions and Hypotheses

4.1 Detecting Intrinsic Power Preferences

I use four methods to detect when behavior in a Monetary Game round is determined by intrinsic power preferences:

- Points Game Criterion: equals 1 if the subject bought power in the analogous Points Game round, and equals 0 otherwise
- Binary Choice Game Criterion (defined for rounds where the subject bought power only): equals 1 if the subject chose the $(12.30, 12.30)$ option in the analogous Binary Choice Game round, and equals 0 otherwise
- Dislikes Power Criterion (defined for rounds where the subject did not buy power only): equals 1 if the subject did *not* choose the $(12.30, 12.30)$ option in the analogous Binary Choice Game round, and equals 0 otherwise
- Unused discretion (defined for rounds where the subject bought power only): discretion purchased in excess of the minimum required to choose the partner's outcome.

The "analogous" Points Game round is the one which shares the same increase and decrease factors as the Monetary Game round in question. The Points Game Criterion makes use of the fact that the Points Game provides no distributional incentive to buy power, meaning any power bought in the Points Game round was motivated by intrinsic power preferences. Because the increase and decrease factors are the same in both games, the discretional incentives to buy power are also the same in both games. Further, since recognition is constant within-subject, recognition incentives to buy power are also the same across the two games. Thus, if behavior in the Points Game is motivated by intrinsic power preferences, then behavior in the Monetary Game is, as well. Meeting the Points Game Criterion is a sufficient but not necessary condition for having positive intrinsic power preferences, since a subject will not buy power if its intrinsic value is lower than the monetary cost of buying power. Note that even in this case, the criterion is informative: failing the Points Game Criterion across all rounds reveals that the intrinsic value of power must be less than 0.25 the value of one's earnings.

For the Binary Choice Game Criterion and Dislikes Power Criterion, the "analogous" Binary Choice Game round is the one with options determined by the subject's decision in the Monetary Game round in question. As a reminder, for spending on power s, change to partner's outcome y, and increase factor u in the Monetary Game, the analogous round of the Binary Choice Game offers two options. Option 1 is always $(12.30, 12.30)$. Option 2 is either $(12.30 - s, 12.30 + y)$ if $s > 0$, or $(11.30, 12.30 + u)$ if $s = 0$. If the subject bought power in the Monetary Game, then the outcome was $(12.30 - s, 12.30 + y)$. Her choice in the Binary Choice Game reveals whether or not she actually prefers that outcome to the no-power outcome: if she chooses $(12.30, 12.30)$, then her choice to buy power must have been motivated by intrinsic power preferences, rather than liking the distribution itself. This is the intuition behind the Binary Choice Game Criterion. Note that failing the criterion does not rule out intrinsic power preferences, nor does meeting the criterion imply that the subject has no distributional preferences. Several motivations could contribute to the value of buying power. Like the Points Game Criterion, meeting the Binary Choice Game Criterion implies positive intrinsic power preferences, but positive intrinsic power preferences do not imply that the subject meets the Binary Choice Game Criterion.

The Dislikes Power Criterion follows a similar logic. If the subject did not buy power in the Monetary Game, yet in the Binary Choice Game she chose an outcome she could have achieved by buying power, she reveals that she had distributional incentives to buy power in the Monetary Game. Thus, not buying power in the Monetary Game implies she had negative power preferences which outweighed her distributional preferences. If a subject has negative power preferences yet stronger distributional preferences, she may buy power; once again, meeting the Dislikes Power Criterion is not a necessary condition for having negative power preferences.

Unused discretion exploits a key feature of the Monetary Game: subjects can choose whether or not to buy more power than they use. When a subject spends s on power, she can choose any outcome for her partner in the range $[-ds, us]$. For the moment, consider the case where $u = d > 0$ and ignore the fact that subjects must choose values divisible by 0.05. If she chooses $-ds$ or us, then she could not choose that outcome by spending less than s, so she has no "unused discretion." If she chooses an outcome on the interior of her budget set, it was possible to have the same outcome for her partner while having higher earnings for herself, so she bought unused discretion. She cannot prefer the distribution she actually chose to this alternative with higher earnings for herself, because that would contradict the assumption that distributional utility is increasing in her own monetary payoff. Thus, the presence of unused discretion is evidence of intrinsic power preferences. She must have gained enough utility from buying unused discretion to offset the reduction in distributional utility from spending earnings.

Now let us consider the case where either $u = 0$ or $d = 0$. If a subject buys power and then does not

change her partner's earnings at all, then she could have done that without buying power, so clearly she has bought unused discretion. However, her choice does technically lie on the edge of her budget set, so we must find a different way to calculate unused discretion.

Now let us account for the fact that a subject must choose her spending and the change to her partner's earnings in increments of \$0.05. Suppose a subject pays s and changes her partner's earnings by y. Then, making use of the fact that $u = d$ in all rounds where $u, d > 0$, we can get the difference between how much she paid and the minimum she needed to pay by rounding $s - \frac{|y|}{\max\{u,d\}}$ down to the nearest 0.05. Call this value $\hat{s}$. From this, we can get the amount of unused discretion by rounding $\hat{s} * \max\{u, d\}$ down to the nearest 0.05. I will use $\hat{y}$ to denote unused discretion.

A skeptical reader may attribute unused discretion to either confusion or error on the part of the subject. To mitigate confusion, all subjects needed to pass a quiz on their comprehension of the rules before each part of the experiment, and this quiz was designed to be difficult to brute-force by guessing. As such, subjects should have understood the costs of their unused discretion before making any decisions. Additionally, unused discretion did not decrease in later rounds of the experiment; we would expect it to decrease with time if subjects were initially confused and then learned with practice, but there is no evidence of that.

Next, consider the possibility of error; for example, subjects may misread the instructions for this round, such that they believe that they can increase their partner's earnings even though they can only decrease them. This would explain why a subject would buy power and then choose not to change her partner's earnings at all when $u = 0$ or $d = 0$, but it does not explain buying unused discretion in rounds where $u, d > 0$. As such, I will pay special attention to rounds where $u, d > 0$ in the results section. Note that once the subject pays for power and is deciding her partner's earnings, the maximum and minimum possible values are displayed on screen, so calculation errors cannot explain picking an interior value. At that point, discretion and recognition are also fixed. The only reason to pick an interior value is that the interior value is best according to her distributional preferences.

4.2 Hypotheses

I propose seven hypotheses. Hypothesis 1 applies the law of demand to power. Hypotheses 2-5 address the conditions under which subjects buy more power. Hypotheses 6-7 concern the measures of intrinsic power preferences.

Hypothesis 1: the amount of power subjects buy is increasing in both the increase factor and decrease factor.

Hypothesis 1 is just the application of the law of demand to power as a good. The increase factor and

decrease factor control the price of power: when these factors are larger, the price of power is lower.

Hypothesis 2a: in the Monetary Game, benevolent subjects buy more positive discretion than negative discretion.

Benevolent subjects always either increase or do not change their partners' earnings, so they are very likely to prefer distributions where their partner's earnings are higher, holding everything else equal.[2] This would mean that positive discretion has distributional value while negative discretion does not. This is one way to rationalize the hypothesis; another is that benevolent subjects may intrinsically prefer positive discretion more than negative discretion. We can determine which rationalization is more plausible by appealing to behavior in the Points Game. If benevolent subjects also buy more positive discretion in the Points Game, then that indicates that they intrinsically prefer positive discretion more than negative discretion, because we can no longer appeal to distributional incentives.

Hypothesis 2b: in the Monetary Game, malevolent subjects buy more negative discretion than positive discretion.

The reasoning behind this hypothesis is analogous to that of hypothesis 2a. If hypotheses 2a and 2b are both supported, we may be able to increase the proportion of benevolent power-holders by structuring positions to have more opportunities to help others than to hurt others. If we instead find that malevolent subjects value all power equally, then it will be difficult to filter them out in this manner.

Hypothesis 3: subjects buy more power in the Monetary Game than in the Points Game.

A Monetary Game round and its analogous Points Game round provide the same incentives regarding discretion and recognition. The only difference is that the Monetary Game allows the subject to change her partner's earnings, while the Points Game does not. Since the subject can still leave her partner's earnings unchanged after buying power in the Monetary Game, the Monetary Game provides a superior set of possible distributions compared to the Points Game. As such, if my theoretical framework encompasses all allures of power and subjects have distributional preferences, then subjects should value power more (and therefore buy more of it) in the Monetary Game than in the Points Game.

Hypothesis 4: subjects in the high information and control treatments buy the same amount of power.

The high information and control treatments differ in self-recognition, since the subject has a 100% chance of being power-holder in the high information treatment compared to a 50% chance in the control treatment. One could plausibly justify both more power and less power being bought in the high information treatment than in the control treatment. In favor of buying more power in the high information treatment: altruistic subjects may feel less obligation to buy power for the sake of helping their partners if they think

[2]This is only "very likely" because it is technically possible that such subjects have no distributional preferences and their consistent increasing is just a coincidence.

there is a 50% chance that their decisions will not matter, anyway. Additionally, to the extent that someone intrinsically enjoys having power, the enjoyment may be diminished if she is uncertain whether she truly has power, so she may spend less in the control treatment (regardless of power type). On the other hand, malevolent and capricious subjects may feel less guilt from buying power and hurting their partners if they can tell themselves there is a 50% chance their decisions will not matter, anyway. Additionally, buying power is costly, so subjects may feel more willing to spend if there is a high chance their choices will not have a cost (regardless of power type). Note that subjects' decisions are only implemented if they are actually the power-holder (e.g., buying power is not actually half as costly in the control treatment), but subjects may view the situations differently, especially when convenient for explaining away a sense of responsibility or guilt.

Hypothesis 5: subjects in the high information and low information treatments buy the same amount of power.

The high information and low information treatments differ in other-recognition, since the partner only learns the game structure and the power-holder's decision in the high information treatment. Once again, there are plausible arguments that spending should be higher in the high information treatment or that it should be lower. In favor of buying more power in the high information treatment, for some, having status as a power-holder in the eyes of others may be desirable, regardless of what one does with it. In addition, benevolent subjects may enjoy getting credit for their good deeds. On the other hand, some people shrink from the spotlight, so to speak, in which case status may be inherently undesirable, regardless of what one does with it. In addition, malevolent or capricious types may feel guilty if their partners learn of their bad deeds.

Hypothesis 6: the Points Game Criterion and Binary Choice Game Criterion are positively correlated.

Both the Points Game Criterion and the Binary Choice Game Criterion aim to detect the presence of positive intrinsic power preferences. Since they aim to detect the same underlying preferences, there should be some agreement between them. I will measure the correlation both at the round-level (whether a subject's behavior in a particular round of the Monetary Game was motivated by intrinsic power preferences) and at the subject-level (how many of a subject's Monetary Game rounds were motivated by intrinsic power preferences). Note that the two criteria fill slightly different niches: meeting the Points Game Criterion means that intrinsic power preferences alone are strong enough to buy power; meeting the Binary Choice Game Criterion means that distributional preferences alone are *not* strong enough to buy power. It is quite possible that a subject's preferences meet one criterion but not the other.

Hypothesis 7: subjects buy unused discretion.

In the Monetary Game, even if a subject is altruistic, so long as she enjoys discretion and/or recognition

strongly enough, she may buy power even when she is only able to decrease her partner's earnings that round (and then not change her partner's earnings at all). However, if a subject buys unused discretion in rounds where she can both increase and decrease her partner's earnings, that implies that she has a satiation point in her distributional preferences. This is not consistent with a wide variety of preferences, including altruism, spitefulness, maximin, and welfare maximization. It may be consistent with inequality aversion if she sets her partner's earnings equal to her own. As such, with heterogeneity in distributional preferences, subjects are overall less likely to buy unused discretion in rounds where they can both increase and decrease their partners' earnings than in rounds where they can only do one or the other.

In the Points Game, I assume that distributional preferences are irrelevant, so subjects should only care about their range of options and not what they actually choose for their partners. As such, the partner's outcome should be chosen essentially randomly, and so, unused discretion should be common. Subjects may be inclined to choose outcomes on the edge of their budget set out of the power of suggestion, simply because the edges of their budget set are displayed on the screen when making the decision. However, I overall expect to see more unused discretion in the Points Game than in the Monetary Game where the partner's outcome has real consequences.

4.3 Exploratory Analyses

In addition to examining power bought for Hypotheses 4-5, I also estimate the the effect of information treatment on the outcome chosen for the power-holder's partner in the Monetary Game and the Points Game. I report the prevalence of intrinsic power preferences according to the Points Game Criterion, Binary Choice Game Criterion, Dislikes Power Criterion, and the frequency of unused discretion.

For each power type, I report the prevalence in the sample, average spending on power, behavior in the Binary Choice Game's common rounds, and average outcome chosen for their partners. I also regress power type on a variety of personal characteristics, including risk aversion, personality type, authoritarianism, demographics, and whether or not the subject had prior lab experience. Given the large number of variables, spurious correlations are very possible, but the results may suggest avenues for future, more targeted research.

I regress both how much subjects spend on power and the outcomes they choose for their partners on personal characteristics. This analysis provides some information about who is likely to self-select into positions of power, but these results should be interpreted cautiously. Not only are there many variables, but also this controlled experiment deliberately removes many of the complexities of real-world contexts. For example, positions of power usually come with opportunities for self-enrichment, there is competition for the position, and people have heterogeneous beliefs about what will happen if someone else occupies the

position.

5 Data

The sample consisted of 193 undergraduate students from the University of Arizona. Subjects were recruited through the University's Economic Science Laboratory website. A summary of subject characteristics is in Table 5. Since there were very few subjects that were American Indian, Black, studying Arts, or studying an unlisted field, I will not make inferences about these categories. Because monthly spending, much like risk aversion, will be treated as a linear rather than categorical variable, I will make inferences about this category despite very few subjects spending more than $3000 per month.

Recall that subjects were randomly assigned to play the Monetary Game before the Points Game or vice versa, and they were also randomly assigned to one of three information treatments. The number of observations per each treatment is reported in Table 6. To control for ordering effects, I include a binary "Points Game first" variable in all regressions, and I separate most non-regression results by which game subjects played first. Additionally, because subject behavior may also change over time, for regressions that treat each round as one observation, I control for where in the order of rounds an observation appears (integer "Round number" variable). I do not find a systematic effect of "Points Game first," (see individual tables in the paper and the Appendix for specific cases of order effects), and "Round number" was never significant.

6 Results

The results are organized into three sections. Section 6.1 addresses hypotheses 1-5. Section 6.2 addresses the prevalence of power types, outcome chosen for partner by power type and information treatment, correlation between outcome chosen and personal characteristics, the prevalence of intrinsic power preferences, and hypotheses 7-8. Section 6.3 addresses spending by power type, Binary Choice Game common round choices by power type, correlation between power type and personal characteristics, and correlation between spending and personal characteristics.

6.1 When People Desire Power

Result 1: subjects buy more power when it is cheaper, regardless of whether it is power to help or hurt.

Hypothesis 1 stated, "the amount of power subjects buy is increasing in both the increase factor and decrease factor." When a subject spends s on power in a round with increase factor u and decrease factor

Table 5: Subject Characteristics

	Mean	SD	Min	Max
Risk aversion[a]	2.65	1.12	0	5
Extraversion[b]	4.40	1.94	0	8
Agreeableness[b]	5.10	1.76	0	8
Conscientiousness[b]	5.46	1.69	1	8
Neuroticism[b]	4.03	2.16	0	8
Openness[b]	5.02	1.74	0	8
Authoritarianism[b]	3.04	2.04	0	8
Religiosity[c]	1.28	1.02	0	3
Prior lab experience[d]	0.64			
Female	0.42			
LGBT+	0.10			
Race:				
White	0.51			
Asian	0.21			
Hispanic	0.17			
Black	0.03			
American Indian	0.01			
Other	0.07			
Political affiliation:				
Democrat	0.29			
Republican	0.26			
Independent	0.31			
Other	0.14			
Field of study:				
Business	0.63			
Science	0.17			
Engineering	0.09			
Social science	0.06			
Arts	0.03			
Other	0.03			
Monthly spending:				
Less than $1000	0.50			
$1000-$2000	0.41			
$2000-$3000	0.07			
$3000-$4000	0.01			
More than $4000	0.01			

Note: This table summarizes the characteristics of the 193 subjects.
[a]Measured on a scale of 0 to 5 (5=most risk averse).
[b]Measured on a scale of 0 to 8 (8=most strongly exhibits trait).
[c]Measured on a scale of 0 to 3 (0=not at all religious; 3=very religious).
[d]Whether or not subjects had participated in similar lab experiments before.

Table 6: Number of Subjects by Treatment

	High information	Low information	Control	Unclassified[a]
Monetary Game first	29	31	30	4
Points Game first	32	31	30	6

Note: This table reports the number of subjects randomly assigned to each treatment.
[a]These subjects are not used in the analysis of information treatment's effect due to a problem during the session.

d, she buys power equal to $us + ds$. I regress power bought on the full interaction of the increase/decrease factors with power type (as well as subject- and round-specific controls; see table notes for details on this and subsequent regressions). I then calculated the marginal effect of the increase and decrease factors for each power type, which are shown in Table 7[3]. Note that, for all regressions, I adjust significance thresholds for multiple hypothesis testing. Hypothesis 1 is clearly supported; for all but selfish types, subjects buy more power when the factors are higher, indicating that power is indeed a good which is responsive to prices. The fact that selfish subjects are not significantly affected by these factors suggests that it was the correct decision to allow them to buy power once; they do seem generally uninterested in power rather than just waiting for low prices.[4]

Table 7: Marginal Effects of the Increase/Decrease Factors on Amount of Power Bought

	(1) Monetary Game	(2) Points Game
Increase factor:		
If benevolent	1.588 (0.138)***	1.263 (0.143)***
If malevolent	0.905 (0.210)***	0.802 (0.213)***
If capricious	1.339 (0.102)***	1.184 (0.108)***
If selfish	0.208 (0.082)	0.256 (0.100)
Decrease factor:		
If benevolent	1.004 (0.079)***	0.819 (0.119)***
If malevolent	1.485 (0.227)***	1.000 (0.278)***
If capricious	1.260 (0.092)***	1.079 (0.108)***
If selfish	0.218 (0.086)	0.200 (0.073)
Observations	2013	2013

Note: This table reports the effect of raising the increase or decrease factor by 1 on power bought by power type, derived from linear regressions. Power is measured in dollars (Monetary Game) or Points (Points Game). Regressions control for the subject's information treatment, the round number, and whether the subject played the Points Game first. Each subject-round constitutes one observation. Unconditional standard errors (clustered at the subject-level) are in parentheses.
Significance level: * $p < 0.00625$, ** $p < 0.003125$, *** $p < 0.000625$.

Result 2: benevolent subjects intrinsically value positive discretion more than negative

[3]The raw regression estimates for hypotheses 1-5 can be found in the Appendix
[4]If I expand the definition of selfish to include subjects who bought power twice, then we do see "selfish" subjects buying more power in rounds with high increase and decrease factors.

discretion.

Result 3: malevolent subjects value negative discretion more than positive discretion for distributional reasons.

Result 3: malevolent subjects value discretion

Hypothesis 2 concerns the relative amounts of positive discretion and negative discretion that subjects bought. To examine this, I regress $us - ds$ (i.e., positive discretion minus negative discretion) on power type. The average vale $us - ds$ by power type is reported in Table 8. I find support for both hypothesis 2a ("In the Monetary Game, benevolent subjects buy more positive discretion than negative discretion") and hypothesis 2b ("In the Monetary Game, malevolent subjects buy more negative discretion than positive discretion"). Interestingly, benevolent subjects also buy significantly more positive discretion than negative discretion in the Points Game, suggesting that they intrinsically value positive discretion more. Meanwhile, malevolent subjects buy both types of discretion in equal proportion in the Points Game, suggesting that the reason they buy more negative discretion in the Monetary Game is that they want to lower their partners' earnings.

Table 8: Average Positive Discretion in Excess of Negative Discretion ($)

	(1) Monetary Game	(2) Points Game
If Monetary Game first:		
If benevolent	0.193 (0.036)***	0.133 (0.036)***
If malevolent	-0.158 (0.049)**	-0.052 (0.053)
If capricious	0.050 (0.030)	0.038 (0.032)
If selfish	0.024 (0.017)	0.026 (0.020)
If Points Game first:		
If benevolent	0.152 (0.031)***	0.124 (0.032)***
If malevolent	-0.198 (0.051)***	-0.065 (0.060)
If capricious	0.009 (0.023)	0.028 (0.029)
If selfish	-0.017 (0.017)	0.014 (0.021)
Observations	2013	2013

Note: This table reports the average difference between positive and negative discretion bought, separated by power type and derived from linear regressions. The difference is measured in dollars (Monetary Game) or Points (Points Game). Regressions control for the subject's information treatment, the round number, and whether the subject played the Points Game first. Each subject-round constitutes one observation. Unconditional standard errors (clustered at the subject-level) are in parentheses.
Significance level: * $p < 0.00625$, ** $p < 0.003125$, *** $p < 0.000625$.

Result 4: benevolent subjects buy the same amount of power in the Points Game as in the Monetary Game.

Hypothesis 3 was, "Subjects buy more power in the Monetary Game than in the Points Game." To test this, I regressed power bought on the full interaction of (1) whether the round was part of the Monetary

Game or Points Game, (2) power type, and (3) whether the subject played the Monetary Game first. The average difference in power bought between the two games is reported in Table 9. First, note that the point estimates suggest an ordering effect: subjects spend less in whichever game they play first. This is more prominent among subjects who played the Monetary Game first; this may be because subjects who played the Points Game without seeing the Monetary Game first take the Points Game more "seriously," as if it were the Monetary Game. Among subjects who played the Monetary Game first, malevolent and capricious types bought significantly more power in the Monetary Game than in the Points Game. Indeed, malevolent types buy \$1.48 less in the Points Game, which is further evidence for Result 3 above. Capricious types also buy less power in the Points Game if it is after the Monetary Game, suggesting that distributional preferences affect how much power they buy. In contrast, benevolent types show no significant difference, which is consistent with Result 2 (for benevolent types, the value of power is intrinsic rather than distributional).

Table 9: Effect of Choosing Points Instead of Earnings on Amount of Power Bought

	If benevolent	If malevolent	If capricious	If selfish
If Monetary Game first	-0.450 (0.231)	-1.478 (0.573)*	-0.472 (0.174)*	0.009 (0.121)
If Points Game first	0.082 (0.176)	0.649 (0.312)	0.223 (0.125)	0.196 (0.221)

Note: This table reports how much higher demand for power is in the Points Game compared to the Monetary Game, derived from a linear regression with 4026 observations. The regression controls for the increase/decrease factors, the subject's information treatment, and the round number. Marginal effects are separated by power type and whether the subject played the Points Game first. Each subject-round constitutes one observation. Unconditional standard errors (clustered at the subject-level) are in parentheses.
Significance level: * $p < 0.0125$, ** $p < 0.00625$, *** $p < 0.00125$.

Result 5: the amount of power bought is not affected by self-recognition and other-recognition.

Hypotheses 4 and 5 pertained to power bought by information treatment. I regressed the amount of power bought on the full interaction of information treatment with power type. The average difference in power bought for the low information and control treatments compared to the high information treatment is shown in Table 10. There is no significant difference between information treatments, supporting both hypothesis 4 ("subjects in the high information and control treatments buy the same amount of power") and hypothesis 5 ("subjects in the high information and low information treatments buy the same amount of power"). Some of the justifications proposed in Section 4.2 for a difference between information treatments only applied to a subset of power types; since we see no difference for all power types, those justifications seem implausible. This is excellent news for experimental economists: the most cost-effective way of running experiments is to structure things like the control treatment, since we can collect data from both the power-holder and the passive player by giving each a 50% chance of being the true power-holder *ex ante*. However, this structure is unusually low in self-recognition compared to real-world positions of power, and it gives the

passive player unusually detailed insight into the power-holder's game. This experiment shows that neither of these factors seem to actually influence how much power subjects buy (and later results will also show no effect on outcome chosen), suggesting that the cost-effective information structure is just as externally valid as other structures.

Table 10: Marginal Effects of Information Treatment on Amount of Power Bought

	(1) Monetary Game	(2) Points Game
Low information:		
If benevolent	-0.066 (0.364)	-0.353 (0.450)
If malevolent	-0.520 (0.525)	0.065 (0.845)
If capricious	0.293 (0.293)	0.681 (0.323)
If selfish	0.123 (0.169)	-0.103 (0.303)
Control:		
If benevolent	-0.022 (0.348)	0.113 (0.430)
If malevolent	0.229 (0.615)	1.255 (0.886)
If capricious	0.356 (0.312)	0.572 (0.356)
If selfish	-0.058 (0.162)	-0.317 (0.250)
Observations	2013	2013

Note: This table reports the difference in power bought among information treatments by power type, derived from linear regressions. The baseline is the high information treatment. Regressions control for the round number, increase factor, decrease factor, and whether the subject played the Points Game first. Each subject-round constitutes one observation. Unconditional standard errors (clustered at the subject-level) are in parentheses.
Significance level: * $p < 0.00625$, ** $p < 0.003125$, *** $p < 0.000625$.

6.2 What People Do With Power

Subjects are classified into power types based on their behavior in the Monetary Game, but I could instead have based their power type on behavior in the Points Game: subjects who change their partners' points at most one time are selfish, non-selfish subjects who only ever increase their partner's earnings are benevolent, etc. The number of subjects who belong to each type according to the two games is cross-tabulated in Table 11. 62% of the sample are consistently classified by the two games; a chi-square test of homogeneity reveals that the distribution of types is different ($p < 0.001$). As such, we must consider which way of classifying power types is better. Because the goal of tracking power types is to identify subjects who would make good power-holders in the real world, I argue that the Monetary Game is a better indicator than the Points Game, since points do not actually affect the well-being of the power-holder's partner.

To determine the relative impact of each power type, we next examine the change they made to their partners' earnings. Figure 1 shows the cumulative distribution function for outcome chosen, separated by game, which game the subject played first, and power type. On average, the effect on partner's earnings was

Table 11: Number of Subjects per Power Type

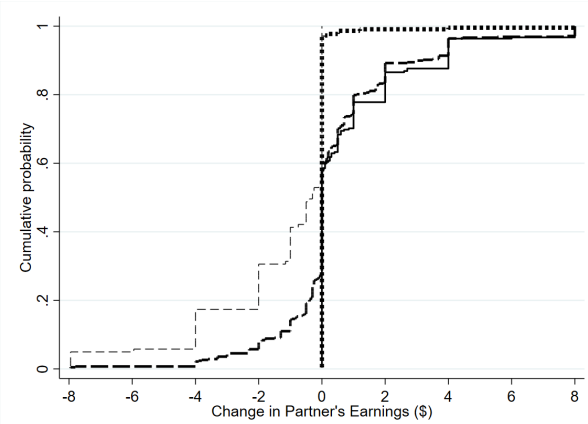
	Points Game				
Monetary Game	Benevolent	Malevolent	Capricious	Selfish	Total
Benevolent	26	3	14	12	55
Malevolent	1	5	9	5	20
Capricious	10	6	50	6	72
Selfish	6	2	0	38	46
Total	43	16	73	61	193

Note: This table reports the number of subjects belonging to each power type. The rows indicate classification by behavior in the Monetary Game; the columns indicate classification by the Points Game.

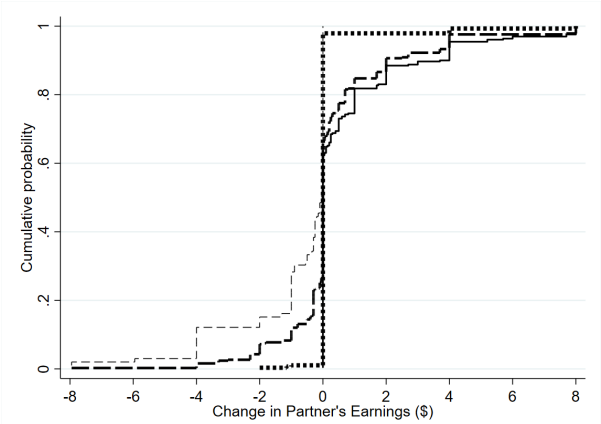
+$0.89 for benevolent types, -$1.11 for malevolent types, +$0.40 for capricious types, and +$0.08 for selfish types. In terms of pure magnitude of change, capricious types ($1.10) are tied for first with malevolent types. Capricious subjects are slightly beneficial to their partners overall, although not to the same degree as benevolent subjects. Because the requirements for being a capricious subject are to have increased one's partner's earnings and decreased one partner's earnings *at least once*, one may wonder how frequently a capricious subject did each, and how their behavior changed over time. Graphs of each capricious subject's outcomes chosen over time are shown in Figure 2. While there is considerable heterogeneity, few subjects show a clear increasing or decreasing trend.

Figure 1: Outcome Chosen for Partner by Power Type

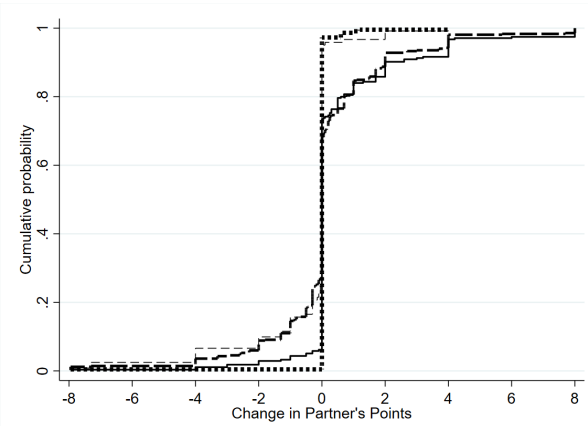
(a) Monetary Game (Monetary Game First)



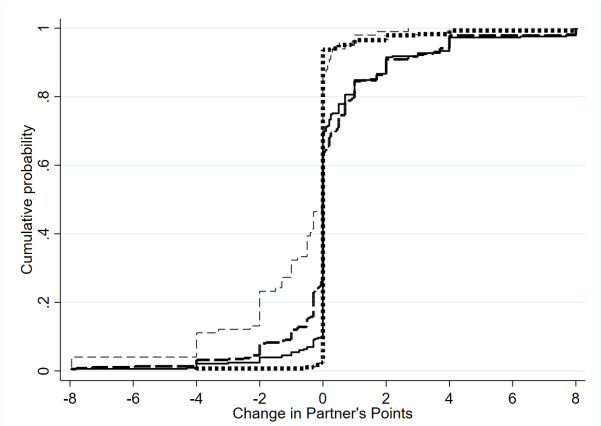
(b) Monetary Game (Points Game First)

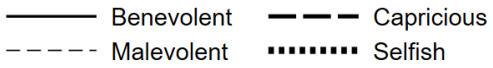


(c) Points Game (Monetary Game First)



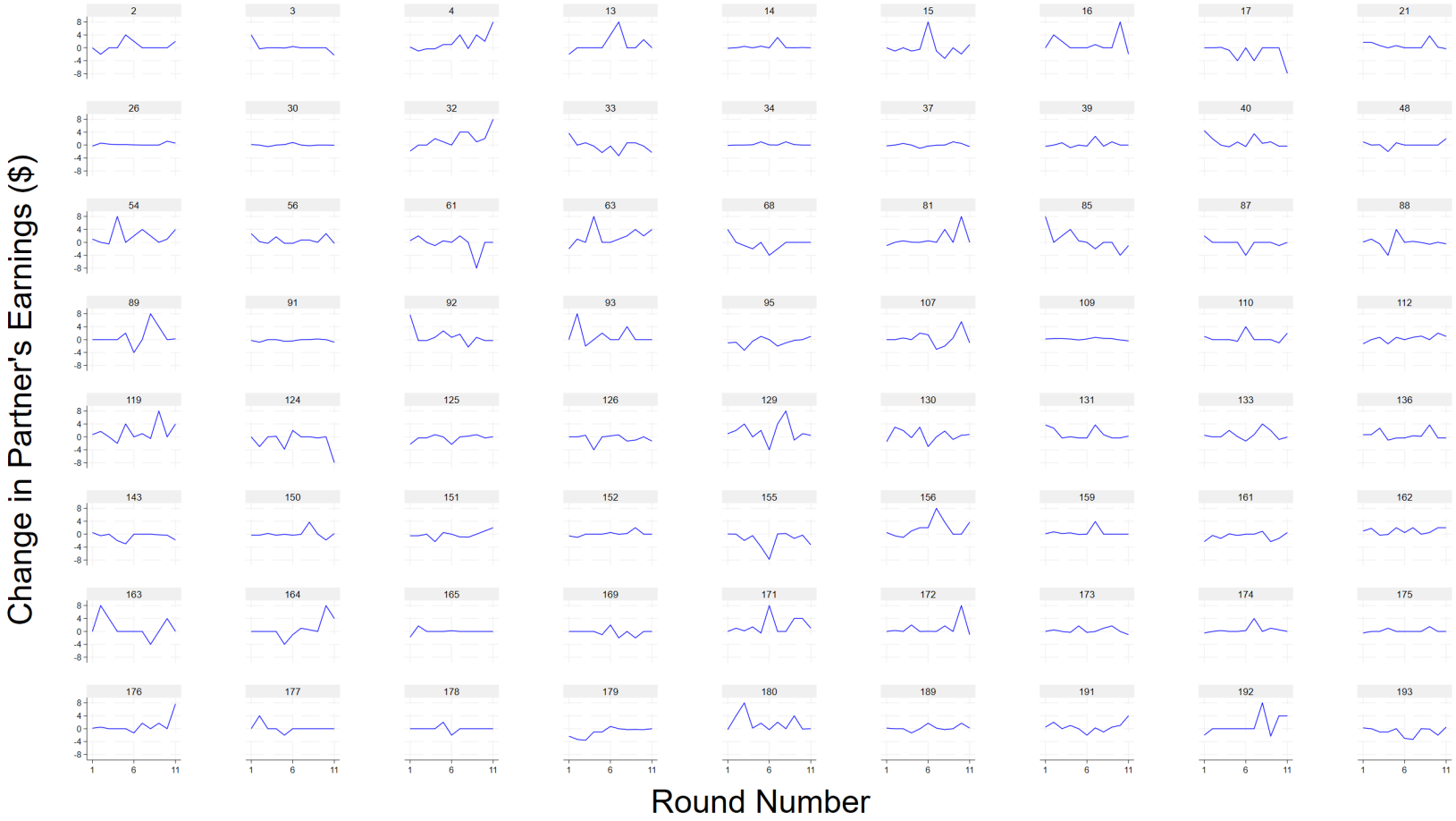
(d) Points Game (Points Game First)





Note: This figure reports the cumulative distribution function of outcomes chosen for the power-holder's partner, measured in dollars (graphs a and b) or points (graphs c and d), for each power type. Each subject-round constitutes one observation. Frequencies are separated by game (graphs a and b for the Monetary Game; c and d for the Points Game), and by whether subjects played the Monetary Game first (a and c) or the Points Game first (b and d).

Figure 2: Capricious Subjects: Monetary Game Outcome Chosen over Time



Note: This figure reports the outcome chosen (measured in dollars) for each round of the Monetary Game, with the rounds in the order that the subject played them. Each graph represents one subject, and only capricious types are shown. The y-axis is the change in the partner's earnings from the no-power outcome ($12.30).

The small average change in the partner's outcome among selfish types reflects the fact that they buy no power in at least 10 out of 11 rounds, which invites the question: how do these numbers change if we restrict our attention to only rounds in which subjects buy power? If we look only at rounds in which the subject bought power, then the average effect on partner's earnings was +$1.94 for benevolent types, -$1.67 for malevolent types, $0.53 for capricious types (average magnitude: $1.49), and +$2.65 for selfish types. This reveals quite a bit of new information. Since benevolent types now have more impact than malevolent or capricious types, the previous result must have been due to malevolent and capricious types buying power more often than benevolent types. Most strikingly, selfish types have the highest impact and were always helpful to their partners. The 35% of selfish types who bought power once may be benevolent people for whom other considerations (e.g., self-interest, disliking discretion) simply discourage them from buying power most of the time.

Result 6: the outcome chosen is not affected by self-recognition and other-recognition.

Next, I estimate the effect of information treatment assignment on the outcome chosen for one's partner. To do this, I regressed outcome chosen on the full interaction of treatment assignment and power type. The difference in outcome chosen relative to the high information treatment is reported in Table 12. As with the amount of power bought, there is no difference whatsoever in outcome chosen among the information treatments, and this holds for all power types. Neither self-recognition nor other-recognition seems to have any effect on behavior in this context.

Table 12: Marginal Effects of Information Treatment on Outcome Chosen

	(1) Monetary Game	(2) Points Game
Low information:		
If benevolent	-0.126 (0.212)	-0.353 (0.450)
If malevolent	0.293 (0.252)	0.065 (0.845)
If capricious	0.065 (0.224)	0.681 (0.323)
If selfish	0.032 (0.080)	-0.103 (0.303)
Control:		
If benevolent	0.047 (0.199)	-0.125 (0.264)
If malevolent	-0.025 (0.278)	-0.455 (0.524)
If capricious	0.169 (0.235)	0.226 (0.265)
If selfish	-0.039 (0.065)	-0.079 (0.145)
Observations	2013	2013

Note: This table reports the difference in outcome chosen for one's partner among information treatments by power type, derived from linear regressions. The baseline is the high information treatment. Regressions control for the round number, increase factor, decrease factor, and whether the subject played the Points Game first. Each subject-round constitutes one observation. Unconditional standard errors (clustered at the subject-level) are in parentheses.
Significance level: * $p < 0.00625$, ** $p < 0.003125$, *** $p < 0.000625$.

So far, we have examined outcome chosen by power type and information treatment. I now regress outcome chosen on the subject's personal characteristics: risk aversion, personality type, authoritarianism, sex, race, political affiliation, field of study, monthly spending (as a proxy for socioeconomic status), whether the subject considers herself part of the LGBT+ community, and whether the subject had prior lab experience. The results are shown in Table 13. The only characteristics which have significant correlation with outcome chosen are being American Indian and being an Arts major, both of which are categories with too few observations to make real inferences. It appears that we cannot predict how a subject will behave from these characteristics, which does not bode well for attracting better power-holders through targeted advertising.

Result 7: according to the Binary Choice Game Criterion, intrinsic power preferences motivated the vast majority of times non-selfish subjects bought power in the Monetary Game.

We now turn to intrinsic power preferences. The proportion of rounds in which subjects met the Points Game Criterion, Binary Choice Game Criterion, or Dislikes Power Criterion, or had any unused discretion, is reported by power type in Table 14. The proportion of subjects who meet the Binary Choice Game Criterion is high for all power types, suggesting that distribution alone is generally not sufficient to convince subjects to buy power. In contrast, the proportion of subjects who meet the Dislikes Power Criterion is quite low. The Dislikes Power Criterion is bound to underestimate the proportion of subjects who dislike power, but the evidence available suggests that such subjects are rare. Unsurprisingly, subjects meet the Points Game Criterion more often if they played the Points Game first rather than after the Monetary Game, regardless of power type. Note that the Points Game Criterion refers to the fraction of all 11 Points Game rounds, while the Binary Choice Game Criterion and unused discretion only take the fraction of Monetary Game rounds in which the subject bought power. Since the Binary Choice Game Criterion focuses only on subjects who were interested enough in power to buy it in the Monetary Game, it is unsurprising that it is met in higher proportion than the Points Game Criterion for all types. Out of the three measures of positive intrinsic power preferences, unused discretion is the least common; it will be addressed in more detail shortly.

Hypothesis 6 was, "the Points Game Criterion and Binary Choice Game Criterion are positively correlated." Let us first examine the correlation for each round individually. Since the Binary Choice Game Criterion is only defined when the subject bought power in the Monetary Game, we will restrict attention to those rounds. The pairwise correlation coefficient is -0.030, which is neither economically nor statistically significant ($p = 0.3359$). Therefore, at the round-level, hypothesis 6 is rejected. Next, let us examine the proportion of rounds in which the subject meets the criteria. Since we are now looking at proportions, we can consider all rounds for the Points Game Criterion and only those where the subject bought power for the Binary Choice Game Criterion. At the subject-level, the correlation coefficient is 0.252, which is significant ($p = 0.0012$) and positive. Thus, at the subject level, hypothesis 6 is supported, although the correlation

Table 13: Estimates of the Relationship between Outcome Chosen and Personal Characteristics

	(1) Monetary Game	(2) Monetary Game	(3) Points Game	(4) Points Game
Risk aversion	-0.018 (0.061)	-0.014 (0.059)	-0.016 (0.069)	-0.012 (0.065)
Extraversion	0.011 (0.028)	0.004 (0.027)	0.005 (0.030)	-0.004 (0.030)
Agreeableness	0.049 (0.036)	0.040 (0.036)	0.046 (0.035)	0.046 (0.037)
Conscientiousness	-0.031 (0.033)	-0.037 (0.033)	-0.002 (0.036)	0.006 (0.037)
Neuroticism	0.003 (0.027)	0.015 (0.028)	0.021 (0.027)	0.002 (0.029)
Openness	0.082 (0.037)	0.082 (0.037)	0.086 (0.032)	0.081 (0.032)
Authoritarianism	0.006 (0.030)	0.017 (0.034)	0.045 (0.030)	0.089 (0.034)
Religiosity		-0.079 (0.065)		-0.105 (0.073)
Prior lab experience		-0.069 (0.132)		-0.017 (0.116)
Female		-0.133 (0.126)		0.190 (0.131)
LGBT+		0.181 (0.198)		0.251 (0.143)
Race:				
Black		-0.022 (0.319)		-0.201 (0.194)
Asian		-0.011 (0.167)		0.087 (0.156)
American Indian		1.287** (0.370)		1.663*** (0.258)
Hispanic		0.056 (0.192)		-0.010 (0.203)
Other		-0.275 (0.239)		-0.310 (0.253)
Political affiliation:				
Republican		0.229 (0.177)		0.103 (0.184)
Independent		0.030 (0.158)		0.187 (0.146)
Other		0.035 (0.166)		0.019 (0.200)
Field of study:				
Science		-0.331 (0.172)		-0.073 (0.146)
Social science		-0.240 (0.284)		-0.204 (0.237)
Arts		0.742 (0.368)		0.987* (0.295)
Engineering		-0.257 (0.158)		0.173 (0.206)
Other		-0.048 (0.259)		0.474 (0.309)
Monthly spending		-0.031 (0.071)		-0.044 (0.079)
Increase factor	0.705*** (0.057)	0.705*** (0.058)	0.538*** (0.056)	0.538*** (0.056)
Decrease factor	-0.172** (0.048)	-0.172** (0.048)	-0.315*** (0.051)	-0.315*** (0.052)
Round number	0.012 (0.012)	0.012 (0.012)	0.016 (0.010)	0.016 (0.010)
Points Game first	0.015 (0.126)	0.048 (0.129)	0.061 (0.128)	0.031 (0.128)
Info treatment:				
Low info	-0.076 (0.139)	-0.040 (0.137)	0.157 (0.142)	0.191 (0.144)
Control	0.053 (0.152)	0.056 (0.144)	0.004 (0.155)	0.017 (0.150)
Constant	-0.829 (0.424)	-0.639 (0.449)	-1.080 (0.446)	-1.184 (0.465)
Observations	2013	2013	2013	2013
Adjusted R^2	0.165	0.185	0.087	0.118
Mean dep. var.	0.309	0.309	0.189	0.189

Note: The coefficients are from linear regressions. Columns (1) and (3) are from regressions which did not include demographic variables; these variables are added in the regressions for columns (2) and (4). Each subject-round constitutes one observation. Standard errors (clustered at the subject-level) are in parentheses. Reference categories: White (race), Democrat (political affiliation), business (field of study), high information (info treatment). Significance level: * $p < 0.002$, ** $p < 0.001$, *** $p < 0.0002$.

Table 14: Frequency of Decisions Motivated by Intrinsic Power Preferences

	Benevolent	Malevolent	Capricious	Selfish
If Monetary Game first:				
Points Game Criterion	0.367	0.397	0.656	0.032
Binary Choice Game Criterion	0.804	0.987	0.956	0.571
Dislikes Power Criterion	0.019	0	0.010	0.032
Unused discretion	0.150	0.184	0.442	0
If Points Game first:				
Points Game Criterion	0.458	0.747	0.714	0.122
Binary Choice Game Criterion	0.743	0.972	0.893	0.778
Dislikes Power Criterion	0.021	0	0.076	0.011
Unused discretion	0.168	0.291	0.431	0.111

Note: This table reports the average fraction of Monetary Game rounds in which a subject met the criteria for intrinsic power preferences, separated by power type and whether the subject played the Monetary Game first. The Binary Choice Game Criterion and unused discretion are calculated as fractions of the Monetary Game rounds in which the subject bought power, where unused discretion equals 1 if the subject had any unused discretion and 0 otherwise. The Dislikes Power Criterion is calculated as a fraction of the Monetary Game rounds in which the subject did not buy power. Each subject constitutes one observation.

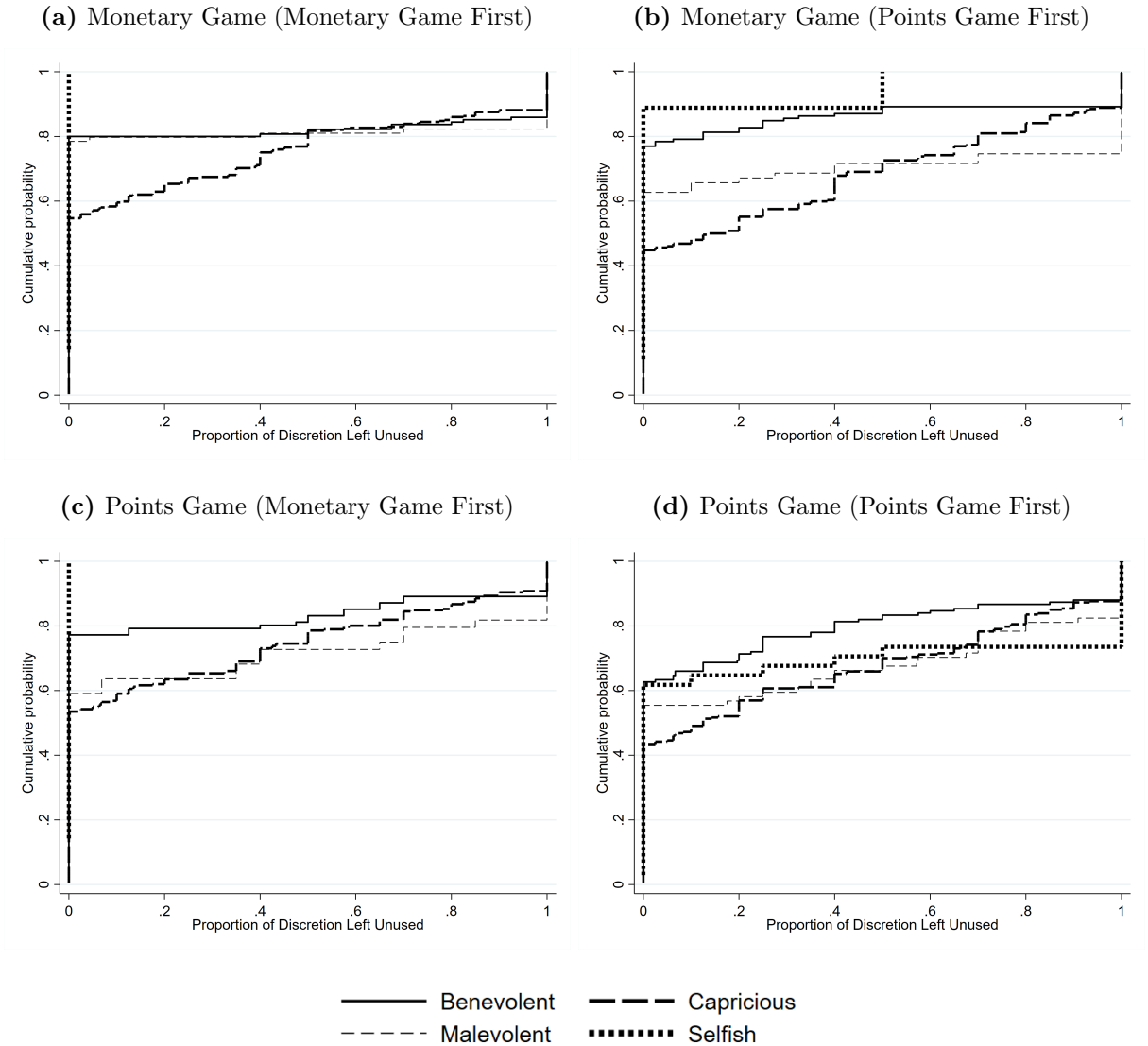
is far from perfect. Recall that subjects will meet the Points Game Criterion if intrinsic power preferences are strong enough to motivate buying power on their own, and subjects will meet the Binary Choice Game Criterion if distributional preferences are not strong enough to motivate buying power on their own. The low correlation between the two criteria shows that simultaneously having both weak distributional preferences and strong intrinsic power preferences is not the norm.

Result 8: subjects buy unused discretion.

Hypothesis 7 was, "subjects buy unused discretion." On average, subjects had \$0.40 of unused discretion in the Monetary Game and \$0.43 in the Points Game. To make behavior between rounds comparable, I will now address the *proportion* of unused discretion, which is unused discretion divided by the biggest change the subject could have made to her partner's outcome given how much she spent (i.e., $\max\{us, ds\}$ rounded down to the nearest 0.05). The cumulative distribution function for the proportion of unused discretion is shown in Figure 3, separated by game, which game the subject played first, and power type. On average, subjects left 24% of their discretion unused in the Monetary Game and 26% unused in the Points Game. Both of these values are significantly greater than zero ($p < 0.001$). However, they are not significantly different from each other, which is not what I predicted (I thought subjects would have more unused discretion in the Points Game since subjects might as well pick the outcome randomly given that points do not affect earnings). I also predicted that subjects would buy more unused discretion in rounds where either the increase factor or decrease factor were zero, since a subject is more likely to have distributional preferences that incentivize changing their partner's earnings in one direction, compared to having distributional preferences that are

maximized on the interior of their choice set (note that these are not mutually exclusive). Indeed, subjects left 34% of their discretion unused when either the increase factor or decrease factor were zero, compared to only 13% when both were positive, so they buy significantly more unused discretion in the former case ($p < 0.001$). Note that 13% is still a significant amount of unused discretion, especially considering the fact that subjects were shown the maximum and minimum earnings they could choose for their partners given what they spent. With this information and the freedom to both increase and decrease their partner's earnings, it is unlikely that the reason for this unused discretion is confusion or calculation errors.

Figure 3: Unused Discretion by Power Type



Note: This figure reports the distribution of unused discretion by power type. Each subject-round constitutes one observation. Frequencies are separated by game (graphs a and b for the Monetary Game; c and d for the Points Game), and by whether subjects played the Monetary Game first (a and c) or the Points Game first (b and d).

To see if power type can be used to predict unused discretion, I run a fractional logit regression of the proportion of unused discretion on power type. The average difference in unused discretion compared to benevolent types is shown in Table 15. The only significant difference is between benevolent and capricious types: capricious types buy about 10% more unused discretion in both games.

Table 15: Marginal Effects of Power Type on Unused Discretion

	(1) Monetary Game	(2) Points Game
Malevolent	0.081 (0.062)	0.126 (0.064)
Capricious	0.110 (0.042)*	0.103 (0.042)*
Selfish	-0.105 (0.059)	0.083 (0.098)
Observations	973	900

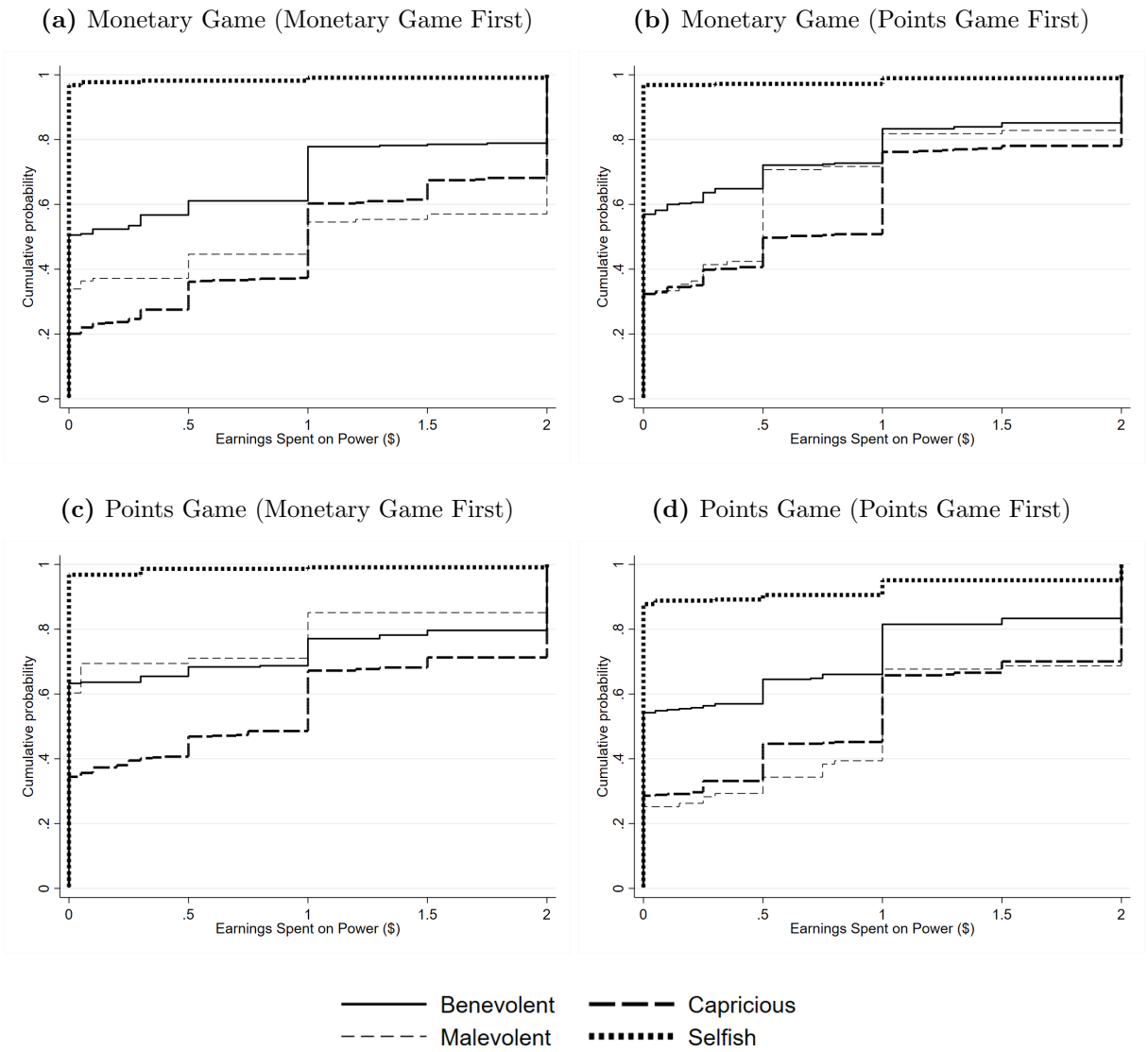
Note: This table reports the difference in the proportion of unused discretion between power types, derived from fractional logit regressions. The baseline category is benevolent. Regressions control for the round number, increase factor, decrease factor, and whether the subject played the Points Game first. Each subject-round constitutes one observation; rounds in which the subject did not buy power are excluded.
Significance level: * $p < 0.016667$, ** $p < 0.008333$, *** $p < 0.001667$.

6.3 Who Desires Power

To begin answering "who desires power," let us first look at power type. Figure 4 shows the cumulative distribution function for amount spent on power. In the Monetary Game, average spending was \$0.56 for benevolent types, \$0.85 for malevolent types, \$0.93 for capricious types, and \$0.03 for selfish types. To test whether spending is significantly different between types, I regress spending on the full interaction of power type with whether the subject played the Monetary Game first. The difference in spending relative to benevolent types is reported in Table 16; capricious types spend significantly more than benevolent types, who in turn spend significantly more than selfish types. If we consider a situation where multiple people must invest resources to increase their chances of winning a competition for power, this result implies that benevolent types will unfortunately be outspent by the less-generous capricious types.

Throughout this paper, I have documented difference in behavior among power types. Is power type a useful concept outside of the two Power Games? And can we predict a person's power type from any personal characteristics? To address the first question, Table 17 summarizes behavior by power type in the 12 Binary Choice Game rounds which were common to all subjects. The first eight decision problems offer the subject a tradeoff between her own earnings and those of her partner (although decision problem 1 makes this choice costless). On average, the altruistic option was chosen 49% of the time by benevolent types, 22% by malevolent types, 37% by capricious types, and 31% by selfish types. These frequencies are significantly different at the 10% level. Thus, benevolent types choose altruistic options the most, while

Figure 4: Earnings Spent on Power by Power Type



Note: This figure reports the cumulative distribution function of earnings (measured in dollars) spent on power for each power type. Each subject-round constitutes one observation. Frequencies are separated by Game (graphs a and b for the Monetary Game; c and d for the Points Game), and by whether subjects played the Monetary Game first (a and c) or the Points Game first (b and d).

selfish/malevolent types choose those options less often. This is indeed reminiscent of their behavior in the Monetary Game. In the other four decision problems, one option offers higher earnings to *both* players than the other (nasty) option. On average, the nasty option was chosen 10% of the time by benevolent types, 11% by malevolent types, 12% by capricious types, and 3% by selfish types. Selfish types are less likely to choose nasty options, but the other differences are not economically significant. Even people who are malevolent when in a position of power are not particularly nasty in this setting.

To answer the question of whether a person's power type can be predicted by their personal characteristics,

Table 16: Marginal Effects of Power Type on Amount Spent on Power ($)

	(1) Monetary Game	(2) Points Game
If Monetary Game first:		
Malevolent	0.339 (0.206)	-0.218 (0.189)
Capricious	0.427 (0.127)**	0.348 (0.152)
Selfish	-0.592 (0.104)***	-0.501 (0.117)***
If Points Game first:		
Malevolent	0.236 (0.125)	0.454 (0.132)***
Capricious	0.364 (0.121)**	0.397 (0.126)**
Selfish	-0.386 (0.081)***	-0.381 (0.124)**
Observations	2013	2013

Note: This table reports the difference in amount spent on power (measured in dollars) between power types, derived from linear regressions. The baseline category is benevolent. Regressions control for the round number, increase factor, decrease factor, and whether the subject played the Points Game first. Each subject-round constitutes one observation. Unconditional standard errors (clustered at the subject-level) are in parentheses.
Significance level: * $p < 0.008333$, ** $p < 0.004167$, *** $p < 0.000833$.

Table 17: Binary Choice Game Common Rounds Choices

(Option 1) vs (Option 2)	Benevolent	Malevolent	Capricious	Selfish
(10.90, 10.90) vs (10.90, 16.60)	0.71	0.35	0.60	0.78
(10.90, 10.90) vs (10.50, 16.60)	0.49	0.15	0.39	0.41
(14.80, 9.60) vs (13.10, 16.60)	0.53	0.15	0.31	0.20
(14.40, 9.50) vs (13.40, 13.40)	0.60	0.35	0.50	0.30
(16.60, 7.80) vs (14.80, 14.80)	0.58	0.30	0.50	0.28
(16.60, 9.40) vs (14.40, 16.60)	0.51	0.15	0.31	0.26
(16.60, 13.90) vs (13.90, 16.60)	0.11	0	0.04	0.02
(16.60, 4.30) vs (10.45, 10.45)	0.42	0.30	0.32	0.20
(7.40, 16.60) vs (4.30, 4.30)	0.16	0.20	0.18	0.02
(16.85, 17.10) vs (16.60, 16.60)	0.27	0.25	0.24	0.11
(10.45, 10.45) vs (4.30, 4.30)	0	0	0.04	0
(16.60, 12.10) vs (12.10, 9.70)	0	0	0.03	0
Observations	55	20	72	46

Note: This table summarizes the fraction of subjects who chose option 2 in the common rounds of Part 4, separated by power type. An option (x, y) offers $\$x$ earnings to the power-holder and $\$y$ earnings to her partner.

I run a multinomial logit regression of power types on personal characteristics with benevolent as the baseline category. The results are shown in Table 18. While the coefficient on LGBT+ is statistically significant here, it is not robust to alternate specifications and so should be disregarded.[5] Therefore, personal characteristics seem to be no indicator of power type, just as they are a no indicator of what outcome a subject will choose for her partner (see Table 13).

Finally, let us examine whether personal characteristics can predict spending on power. To do this, I regress spending on personal characteristics; the results are shown in Table 19. The only characteristic with significant effect is the personality trait of openness: an increase of one point on the eight-point openness scale tends to increase spending in the Monetary Game by about 8 cents, and it looks like there is a slight positive effect in the Points Game, as well. Agreeableness seems to have a similar effect as openness in both games. Note that some prior literature has found a positive correlation between agreeableness and altruism (Baumert et al., 2014; Becker et al., 2012; Ben-Ner, Kramer, and Levy, 2008; Koole et al., 2001; LePine and Van Dyne, 2001), but I only find an effect on spending, *not* outcome chosen. Other than these two personality traits, most personal characteristics seem generally unhelpful for predicting behavior in situations with power. This lack of correlation is an interesting result in its own right.

7 Conclusion

In this paper, I investigate the pursuit of power over the outcomes of others along three dimensions: "who desires power," "when do they desire it," and "what do they do with it?" I find that capricious people are the most common type of power-holder, and they spend more money on buying power than other types. This includes outspending benevolent people, which suggests that the voluntary nature of entering many positions of power may result in less-than-ideal power-holders. Granted, in this experiment, the alternative to seeking power was deterministic and resulted in equal payoffs; in real-world settings, benevolent people may seek out power more often than seen here, motivated to prevent malicious actors from filling those influential roles in their stead.

I find a plethora of evidence that power is valued for more than the monetary outcomes it enables. My results support this conclusion in multiple ways. First, subjects often bought more power than was necessary to bring about the outcome they actually chose for their partner. This implies that people value "unused discretion": the presence of options that they do not take. Indeed, once subjects are presented with the choice between either the outcome they chose in the Monetary Game or exactly one other possible outcome in that game, many subjects who bought power indicate that they prefer the outcome they would

[5]The alternate specification results are available from the author upon request.

Table 18: Estimates of the Relationship between Power Type and Personal Characteristics

	Capricious	Malevolent	Selfish
Risk aversion	1.294 (0.252)	1.033 (0.364)	1.135 (0.236)
Extraversion	1.176 (0.148)	1.213 (0.155)	1.129 (0.179)
Agreeableness	0.815 (0.122)	0.913 (0.207)	0.658 (0.110)
Conscientiousness	1.179 (0.165)	1.364 (0.289)	1.193 (0.204)
Neuroticism	0.881 (0.121)	0.938 (0.121)	0.593 (0.100)
Openness	1.055 (0.166)	0.717 (0.164)	0.763 (0.153)
Authoritarianism	1.124 (0.149)	1.151 (0.177)	1.059 (0.165)
Religiosity	0.824 (0.202)	1.154 (0.371)	0.557 (0.195)
Prior lab experience	0.750 (0.362)	0.887 (0.535)	1.223 (0.776)
Female	2.274 (1.317)	2.437 (1.627)	5.113 (3.600)
LGBT+	0.034*** (0.030)	1.131 (1.183)	0.211 (0.243)
Race:			
Black[a]	7.058 (8.949)	-	1.716 (2.719)
Asian	1.410 (0.905)	1.156 (1.041)	2.407 (2.021)
American Indian[b]	-	-	-
Hispanic	0.835 (0.545)	0.564 (0.609)	0.211 (0.189)
Other	0.757 (0.678)	1.465 (1.502)	0.993 (0.999)
Political affiliation:			
Republican	0.379 (0.236)	0.857 (0.760)	0.190 (0.145)
Independent	0.908 (0.552)	1.563 (1.221)	1.011 (0.679)
Other	3.606 (2.944)	0.878 (1.348)	2.843 (2.564)
Field of study:			
Science	0.568 (0.371)	4.378 (3.635)	0.578 (0.442)
Social science	1.880 (2.204)	5.636 (8.285)	3.081 (3.548)
Arts[c]	0.043 (0.091)	-	-
Engineering	0.605 (0.467)	0.565 (0.620)	1.597 (1.377)
Other	0.082 (0.086)	1.018 (1.431)	0.137 (0.201)
Monthly spending	0.494 (0.154)	0.629 (0.305)	0.800 (0.289)
Points Game first	0.651 (0.316)	0.543 (0.427)	1.248 (0.698)
Information treatment:			
Low information	0.640 (0.390)	3.810 (3.799)	0.851 (0.582)
Control	0.279 (0.151)	1.710 (1.829)	0.164 (0.138)
Constant	2.887 (5.776)	0.063 (0.139)	94.992 (219.954)
Observations		183	

Note: The coefficients are from a multinomial logit regression. The baseline category is benevolent. Each subject constitutes one observation. Standard errors are in parentheses. Reference categories: White (race), Democrat (political affiliation), business (field of study), high information (information treatment).
Significance level: * $p < 0.001333$, ** $p < 0.000667$, *** $p < 0.000133$.
[a]No Black subjects were malevolent.
[b]All American Indian subjects were benevolent.
[c]None of the subjects studying arts were malevolent or selfish.

Table 19: Estimates of the Relationship between Spending on Power and Personal Characteristics

	(1) Monetary Game	(2) Monetary Game	(3) Points Game	(4) Points Game
Risk aversion	0.043 (0.035)	0.042 (0.033)	-0.020 (0.041)	-0.018 (0.039)
Extraversion	0.022 (0.018)	0.023 (0.017)	0.018 (0.019)	0.011 (0.020)
Agreeableness	0.070 (0.023)	0.056 (0.021)	0.057 (0.026)	0.043 (0.024)
Conscientiousness	0.032 (0.022)	0.042 (0.024)	0.006 (0.024)	0.008 (0.025)
Neuroticism	0.022 (0.017)	0.030 (0.018)	0.024 (0.017)	0.026 (0.021)
Openness	0.069** (0.020)	0.079** (0.022)	0.041 (0.022)	0.046 (0.025)
Authoritarianism	0.055 (0.018)	0.045 (0.021)	0.061 (0.021)	0.049 (0.025)
Religiosity		-0.021 (0.047)		-0.010 (0.053)
Prior lab experience		-0.005 (0.079)		-0.084 (0.090)
Female		-0.045 (0.086)		-0.019 (0.101)
LGBT+		-0.173 (0.105)		-0.143 (0.140)
Race:				
Black		0.109 (0.196)		-0.035 (0.134)
Asian		0.095 (0.114)		0.066 (0.126)
American Indian		1.422*** (0.216)		1.070*** (0.126)
Hispanic		0.270 (0.093)		0.286 (0.126)
Other		0.004 (0.146)		0.137 (0.168)
Political affiliation:				
Republican		0.193 (0.114)		0.110 (0.126)
Independent		-0.043 (0.088)		-0.085 (0.104)
Other		0.149 (0.142)		0.053 (0.156)
Field of study:				
Science		-0.063 (0.094)		-0.139 (0.101)
Social science		0.006 (0.195)		-0.029 (0.184)
Arts		0.393 (0.139)		0.330 (0.210)
Engineering		-0.305 (0.110)		-0.230 (0.167)
Other		-0.028 (0.156)		0.290 (0.237)
Monthly spending		-0.097 (0.043)		-0.022 (0.052)
Increase factor	0.122*** (0.020)	0.122*** (0.020)	0.105*** (0.018)	0.105*** (0.018)
Decrease factor	0.034 (0.017)	0.034 (0.017)	0.005 (0.018)	0.005 (0.018)
Round number	0.002 (0.004)	0.002 (0.004)	-0.003 (0.004)	-0.003 (0.004)
Points Game first	-0.241* (0.074)	-0.188 (0.074)	0.075 (0.085)	0.085 (0.086)
Info treatment:				
Low info	0.081 (0.092)	0.136 (0.089)	0.060 (0.105)	0.105 (0.109)
Control	0.107 (0.089)	0.133 (0.091)	0.082 (0.101)	0.091 (0.106)
Constant	-0.867* (0.274)	-0.950 (0.344)	-0.438 (0.282)	-0.334 (0.354)
Observations	2013	2013	2013	2013
Adjusted R^2	0.149	0.218	0.072	0.118
Mean dep. var.	0.596	0.596	0.586	0.586

Note: The coefficients are from linear regressions. Each subject-round constitutes one observation. Columns (1) and (3) are from regressions which did not include demographic variables; these variables are added in the regressions for columns (2) and (4). Standard errors (clustered at the subject-level) are in parentheses. Reference categories: White (race), Democrat (political affiliation), business (field of study), high information (info treatment). Significance level: * $p < 0.002$, ** $p < 0.001$, *** $p < 0.0002$.

have gotten by not buying power, implying that gaining the additional, unused options was a key part of the appeal of buying power. Additionally, many subjects buy about as much power when it grants control over meaningless points as when it grants control over experimental earnings. Overall, this experiment has provided much evidence that it is *power itself* which people find alluring.

I also find that people are sensitive to the price of power. Regardless of whether a subject is benevolent, malevolent, or capricious, she will buy more power as the prices of increasing her partner's outcome and decreasing it are lowered.

I find that power-holders do not seem to care about recognition as a power-holder, either by themselves or others. They behave similarly regardless of whether the person whose outcome they control learns of their decision or not, so recognition by others is unimportant. They also behave the same when they are confident that they are the power-holder as when they are just as likely as their partner to be the power-holder, so self-recognition is unimportant. While this makes running experiments considerably cheaper and easier, it is also a surprising result. Many real-world positions of power come with fame and status, which are usually thought to be pertinent aspects of the job. To the extent that this is true, such fame and status cannot, apparently, be boiled down to the fact that oneself or others know that one has power. In this experiment, whether a subject had power was randomly determined, so perhaps fame and status of real positions come from the ways that one earns the position. Future studies may investigate this further.

Future studies could also focus on whether preferences for discretion (i.e., having options over other people's outcomes) are correlated with preferences for autonomy (i.e., having options over one's *own* outcomes). This paper deliberately removed concerns such as other people occupying the position of power if one does not seize it oneself; behavior may differ substantially once this is a possibility, especially if there is costly competition for the position.

The primary result from this paper is that the intrinsic allure of power is real and widespread. Power is not merely a means to an end. Even when power does not benefit oneself, and even when it does not affect the monetary state of others, people will often sacrifice their own earnings to attain it. Since this experiment stripped away self-interested and distributional motives, what remains is the intrinsic value of power: the acquisition of options which one may or may not use. Given how common intrinsic preferences for power are, economists are likely to mispredict behavior if they look solely at the outcomes of decisions and ignore the options that people have along the way.

References

Ambuehl, S., Bernheim, B. D., & Ockenfels, A. (2021). What motivates paternalism? An experimental study. *American Economic Review*, *111*(3), 787–830. https://doi.org/10.1257/aer.20191039

Andreoni, J., & Miller, J. (2002). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737–53. https://doi.org/10.1111/1468-0262.00302

Bartling, B., Cappelen, A. W., Hermes, H., Skivenes, M., & Tungodden, B. (2023). *Free to fail? Paternalistic preferences in the United States* (Discussion Paper No. 16151). IZA - Institute of Labor. https://www.jstor.org/stable/resrep67192

Bartling, B., Fehr, E., & Herz, H. (2014). The intrinsic value of decision rights. *Econometrica*, *82*(6), 2005–39. https://doi.org/10.3982/ECTA11573

Baumert, A., Schlösser, T., & Schmitt, M. (2014). A performance-based assessment of fairness and altruism. *European Journal of Psychological Assessment*, *30*(3), 178–192. https://doi.org/10.1027/1015-5759/a000183

Becker, A., Deckers, T., Dohmen, T., Falk, A., & Kosse, F. (2012). The relationship between economic preferences and psychological personality measures. *Annual Review of Economics*, *4*, 453–78. https://doi.org/10.1146/annurev-economics-080511-110922

Ben-Ner, A., & Kramer, A. (2011). Personality and altruism in the dictator game: Relationship to giving to kin, collaborators, competitors, and neutrals. *Personality and Individual Differences*, *51*, 216–221. https://doi.org/10.1016/j.paid.2010.04.024

Ben-Ner, A., Kramer, A., & Levy, O. (2008). Economic and hypothetical dictator game experiments: Incentive effects at the individual level. *The Journal of Socio-Economics*, *37*, 1775–1784. https://doi.org/10.1016/j.socec.2007.11.004

Bolton, G. E., & Ockenfels, A. (2000). Erc: A theory of equity, reciprocity, and competition. *American Economic Review*, *90*(1), 166–93. https://doi.org/10.1257/aer.90.1.166

Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, *117*(3), 817–69. https://doi.org/10.1162/003355302760193904

Dana, Cain, & Dawes. (2006). What you don't know won't hurt me: Costly (but quiet) exist in dictator games. *Organizational Behavior and Human Decision Processes*, *100*(2), 193–201. https://doi.org/10.1016/j.obhdp.2005.10.001

Engelhardt, A. M., Feldman, S., & Hetherington, M. J. (2023). Advancing the measurement of authoritarianism. *Political Behavior*, *45*, 537–60. https://doi.org/10.1007/s11109-021-09718-6

Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review*, *103*(4), 1325–59. https://doi.org/10.1257/aer.103.4.1325

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, *114*(3), 817–868. https://doi.org/10.1162/003355399556151

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–78. https://doi.org/10.1007/s10683-006-9159-4

Freundt, J., Herz, H., & Kopp, L. (2023). *Intrinsic preferences for choice autonomy* (Working Paper No. 10342). CESifo. https://doi.org/10.2307/2525186

García-Gallego, A., Georgantzis, N., & Ruiz-Martos, M. J. (2019). The heaven dictator game: Costless taking or giving. *Journal of Behavioral and Experimental Economics*, *82*, Article 101449. https://doi.org/10.1016/j.socec.2019.101449

Grossmann, M. R. P. (2024). *Knowledge and freedom: Evidence on the relationship between information and paternalism*. arXiv.org. https://arxiv.org/abs/2410.20970

Grossmann, M. R. P. (2025). *Paternalism and deliberation: An experiment on making formal rules*. arXiv.org. https://arxiv.org/abs/2501.00863

Koole, S. L., Jager, W., van den Berg, A. E., Vlek, C. A. J., & Hofstee, W. K. B. (2001). On the social nature of personality: Effects of extraversion, agreeableness, and feedback about collective resource use on cooperation in a resource dilemma. *Personality and Social Psychology Bulletin*, *27*(3), 289–301. https://doi.org/10.1177/0146167201273003

LePine, J. A., & Van Dyne, L. (2001). Contextual performance: Evidence of differential relationships with Big Five personality characteristics and cognitive ability. *Journal of Applied Psychology*, *86*(2), 326–36. https://doi.org/10.1037/0021-9010.86.2.326

Meemann, C. (2023). *On the economic value of decision rights: An experimental test* [Doctoral dissertation, Helmut-Schmidt-Universität/Universität der Bundeswehr Hamburg].

Owens, D., Grossman, Z., & Fackler, R. (2014). The control premium: A preference for payoff autonomy. *American Economic Journal: Microeconomics*, *6*(4), 138–61. https://doi.org/10.1257/mic.6.4.138

Pikulina, E. S., & Tergiman, C. (2020). Preferences for power. *Journal of Public Economics*, *185*, Article 104173. https://doi.org/10.1016/j.jpubeco.2020.104173

Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, *41*, 203–12. https://doi.org/10.1016/j.jrp.2006.02.001

Sloof, R., & von Siemens, F. A. (2017). Illusion of control and the pursuit of authority. *American Economic Journal: Microeconomics*, *20*, 556–73. https://doi.org/10.1007/s10683-016-9499-7

Visser, M. S., & Roelofs, M. R. (2011). Heterogeneous preferences for altruism: Gender and personality, social status, giving and taking. *Experimental Economics*, *14*, 490–506. https://doi.org/10.1007/s10683-011-9278-4