# Altruistic Punishment in Elections

Jason A. Aimone
Luigi Butera
Thomas Stratmann

# Altruistic Punishment in Elections

## Abstract

Altruistic punishment is a fundamental driver for cooperation in human interactions. In this paper, we expand our understanding of this form of costly punishment to help explain a puzzle of voting behavior: why do people who are indifferent between two potential policy outcomes of an election participate in large-scale elections when voting is costly? Using a simple voting experiment, we show that many voters are willing to engage in voting as a form of punishment, even when voting is costly and the voter has no monetary stake in the election outcome. In our sample, we observe that at least fourteen percent of individuals are willing to incur a cost and vote against candidates who broke their electoral promises, even when they have no pecuniary interest in the election outcome.

*Jason A. Aimone*
*Baylor University*
*jason.aimone@gmail.com*

*Luigi Butera*
*George Mason University*
*lbutera2@gmail.com*

*Thomas Stratmann*
*Department of Economics*
*George Mason University*
*4400 University Drive*
*USA - 22030 Fairfax VA*
*tstratma@gmu.edu*

August 8, 2014

# 1 Introduction

Since the seminal works of Downs (1957) and Riker and Ordenshook (1968), scholarly research has sought to explain the existence of the massive levels of voting observed in general elections. That is because studies based on rational choice models tend to predict vanishing voter turnout in large-scale elections (Palfrey and Rosenthal 1983; Ledyard 1984; Palfrey and Rosenthal 1985). Fiorina coined the failure of these models that predict voter turnout as "the paradox that ate rational choice."[1] Though scholars have studied voting in elections for at least over half a century, Shughart (2004, p.329) suggests, "the act of voting [...] has eluded explanation on narrow self-interest grounds. The resolution of the "paradox of voting" remains one of the most important challenges facing public choice scholars."

Here, we show that voters who are otherwise indifferent between policies offered by electoral candidates are willing to incur the costs of voting when casting a ballot allows them to vote against candidates who have broken a campaign promise. They do so, even when not personally financially harmed by the broken campaign promise.

In a laboratory experiment, we find that some decisions to cast a costly ballot are due to altruistic punishment voting. We show that this type of voting may explain the behavior of between ten to twenty percent of indifferent voters' decisions to cast a ballot, rather than to abstain.

The classic model by Downs (1957) assumes that voting is instrumental, so that voters cast a ballot to increase the probability that their preferred candidate wins. Much of the subsequent work in the economics and political science literatures have built on this original work, which is often referred to as a pivotal voters model (Fiorina 1976).

A parallel line of voting research has sought to explain voting by civic duty, habit, peer pressure, and other motives. This type of voting, due to a sense of civic duty, or because it increases a person's utility, has been referred to as expressive voting (Brennan and Buchanan 1984; Brennan and Lomasky 1997; Posner 2002; Tyran 2004; Feddersen and Sandroni 2006; Fowler 2007; Hillman 2010). Thus, even individuals whose material well-being is in principle unaffected by an election may nevertheless vote, for instance because of a sense of ethical obligation (Harsanyi 1977; Blais 2000; Feddersen and Sandroni 2006) or social pressure (Gerber, Green and Larimer 2008). Alternatively, voters might cast a ballot to signal to both themselves and others regarding their position on moral issues (Lyndbeck, Nyberg and Weibull 1999; Tyran 2004), or simply to cheer or express sympathy for a certain political view (see i.e. Brennan and Hamlin 1998).[2]

More generally, individuals may be willing to vote to express their preferences over social norms, even when they are not materially affected by elections' outcome (see i.e. Posner 2002). Recent work by Della Vigna et al. (2014) shows in a field experiment that social image is a significant determinant of voter turnout.

---

[1]See, for example, Feddersen (2004).
[2]For a review of this literature see Geys (2006).

While scholars have proposed many theories to explain why people vote, altruistic punishment has not been considered as an explanation for voter turnout in large-scale elections. Possibly, this is because such a motivation for voting requires no preferences over policy platforms or intrinsic preference for voting. Nonetheless, understanding whether altruistic punishment plays a systematic role in candidate elections represents a fundamental empirical question.

Estimating the role that punishment plays in elections using natural occurring data presents several challenges, such as for example, what defines a broken promise. We therefore adopt an experimental laboratory methodology to provide us with rigorous identification strategy. In our experiment we study whether potential voters are willing to engage in costly voting when casting a ballot can serve as an altruistic punishment device for an incumbent who previously broke a campaign promise. We study whether voters engage in this costly voting activity even when this betrayal does not affect these voters directly, and when this betrayal does not generate a social deadweight loss.[3]

Altruistic punishment can promote cooperation and deter norm violations, anti-social, and free-riding behaviors (e.g.Fehr and Fischbacher 2004); or Boyd, Gintis, Bowles and Richerson 2003). For example, in situations where individuals are asked to contribute personal resources to a public good, punishment mechanisms represent an effective solution to free-ride problems (violations of the norm of strong-reciprocity, see Fehr and Fischbacher 2004. In these experiments, punishment increases cooperation (see, for example, Fehr and Fischbacher 2003). Punishment in public goods games is often called ?Altruistic Punishment? (see for example Fehr and Gaechter 2002; Fowler 2005) because such punishment is costly, provides benefits to others who did not undertake the punishment cost, and punishment is done by 3rd parties who are unaffected by free-rider actions. Moreover, there is no repeated interactions so a punisher cannot benefit from changes in the actions of the punished. Thus, such punishment cannot be explained by direct reciprocity. Interestingly, past research (Fowler et al., 2005 and Johnson et al. 2009) also shows a strong connection between egalitarian motives, like inequality aversion, and the willingness to engage in altruistic punishment. Since our goal is to identify altruistic punishment of norm violations as a driving motivation of voting decisions, we limit our definition here to punishment from individuals with no future personal stake in election outcomes, and where we can confidently eliminate egalitarian concerns as a motivator for punishment. In other words we design an experiment to identify altruistic punishment as a motivation for engaging in costly voting.

In our design, we exogenously assign to participants in our experiment preferences over political platforms. Seven voters participate in two sequential elections where candidates make promises about which platform (A or B) to implement,[4] and where casting a vote has a monetary cost. One partisan voter favors platform A and one partisan voter favors

---

[3]For instance, no social deadweight loss occurs because the only consequence of betrayal is wealth redistribution from one group of voters to another group of voters.

[4]Platforms prescribe how a fixed amount of money will be distributed among voters.

platform B, and both favor their preferred platform with equal intensity. The remaining five voters receive the same payoff regardless of which platform is implemented. Voters' payoffs are common knowledge among all voters and candidates, implying that indifferent voters know that the payoffs for the two partisan voters are symmetric with respect to the platforms.

This design implies that indifferent voters have no a priori reason to favor one candidate over the other. Another consequence of this design is that voters have no monetary incentive to punish the incumbent in the second election when that incumbent has not implemented the platform that he promised prior to his first election. This is because the betrayal results in a redistribution of wealth from one partisan voter to another partisan voter, whose positions are symmetric. Economic studies on trust have shown that there are a great deal of emotions related to the act of betrayal, which affect people above and beyond pecuniary gains and losses (Bohnet et al.,2008; or Aimone and Houser, 2013). Such norms against betrayal, of the promise in this case, may be the broken norm that stimulates altruistic punishment voting.[5] Further, Fehr and Gaechter (2002) connect emotional responses as strongly related to altruistic punishment.

Our results provide the first unambiguous evidence of the use of voting as a form of altruistic punishment. Further, we provide the first estimate of the lower bound of fraction of voters whose behavior is consistent with altruistic punishment. We conservatively estimate that 12 percent of politically indifferent voters have preferences that lead them to engage in costly voting as a form of punishment. Our results suggest that citizens who are indifferent between political platforms may have a significant role in swinging elections in disfavor of untrustworthy politicians. Further, the presence of punishment of intentions in elections (since outcomes are irrelevant to indifferent voters) supports the notion that retrospective voting represents an important element for electoral outcomes (see i.e. Fiorina 1981; Svoboda 1995; Tufte 1975; Kinder and Kiewiet 1979; Abramowitz 1988).

The next section of this paper describes the theoretical model, followed by the experimental design. We then show the results and the final section concludes.

## 2  Theoretical Framework

Our simple theoretical model shows how altruistic punishment and inequity aversion determines voting behavior in elections. Here, we draw on the model proposed by Della Vigna, List, Malmendier and Rao (2014).

Suppose two sequential elections take place in time $t = 1, 2$. In election 1 two candidates, $x$ and $y$, run for election. Candidates offer one of two platforms ($A$ and $B$) which

---

[5]Norm formation is a large and important line of research that we do not directly address in this paper. For more research on this topic in a context related to altruistic punishment see Fehr and Fischbacher 2004. For research that that more generally analyzes how norms facilitate the emergence of trade see Kimbrough et al. 2006.

differ only in the way wealth is distributed across voters. Before the election takes place, candidates make non-binding promises regarding which platform they will implement if elected.

We assume that the two platforms are symmetric, in the sense that the total amount of wealth distributed is the same, but under $A$, $n$ voters receive a higher payoff than $m$ voters, while the opposite is true under $B$.[6] The winner of the election can either break the promise he made prior to the election and implement the opposite platform ($b = 1$), or keep the promise and implement the proposed platform ($b = 0$).

In $t = 2$ a second election takes place, where the incumbent from the first election runs against a new candidate $z$. This incumbent, who has already committed to a platform in $t = 1$, will maintain his choice also in $t = 2$, should he be reelected. This means that in $t = 2$ the incumbent does not make new campaign promises. The new candidate $z$ instead can choose either platform $A$ or $B$, and make non-binding campaign promises before the second election.

In the second election, the platforms $A$ and $B$ available to candidates have the identical payoff structure as the platforms in the first election.

There are three types of voters: $V_A$, $V_B$, and $V_I$. There are $n$ voters $V_A$, which are voters who receive higher payoffs if platform $A$ is implemented (A-partisans). Symmetrically, there are $m$ voters $V_B$ who receive higher payoffs when platform $B$ is implemented (B-partisans). Further, voters $V_I$ are indifferent between the two platforms, that is, their monetary payoff is the same regardless of the outcome of the election.

We assume that voting depends on four factors. These are the expected benefit from being pivotal, warm-glow, the cost of voting, and social preferences.

Defining $i = \{V_A, V_B, V_I\}$, a generic type $i$ voter will vote for candidate $k$ if the net expected utility from voting is positive:

$$\sum_{t=1}^{2}(p \cdot \nu_{it} + g_{it} - c_{it}) + b \cdot \text{MAX}(D_v; D_{nv} - C_{nv}) + e \cdot \text{MAX}(p \cdot \sum_{t=1}^{2}|\nu_{At} - \nu_{Bt}|; 0) \geq 0 \quad (1)$$

Similar to Della Vigna et al. (2014), the first sum in (1) represents the standard model of voting. The first term of the sum, $p \cdot \nu_{it}$, represents the expected benefit from being pivotal, with $p$ being the probability of being pivotal in election $t$, and $\nu_{it}$ the benefit for voter type $i$ from pivotality. The term $g_{it}$ captures the warm-glow[7] from voting in election $t$. We assume warm-glow is time-invariant, that is, $g_{i1} = g_{i2}$. The third term of the sum, $-c_{it}$, is the cost of voting.

The core part of our model is the term $b \cdot \text{MAX}(D_v; D_{nv} - C_{nv}) + e \cdot \text{MAX}(p \cdot \sum_{t=1}^{2}|\nu_{At} - \nu_{Bt}|; 0)$, which represents the "social preferences" component of utility.

---

[6]We further assume that $n = m$.

[7]This may include warm-glow from exercising civic duty.

The first social preference component derives from individuals' tolerance for betrayal $b$, and related costs and benefits for punishing betrayal when perpetrated. Suppose the winning candidate breaks his promise ($b = 1$) in $t = 1$. If a voter acts upon betrayal (e.g. punish the incumbent by voting against the incumbent), the voter receives utility $D_v$.[8] If the voter instead does not cast a ballot against the winner (that is, either he abstains or votes for the opposition candidate), the voter derives utility $D_{nv} - C_{nv}$. Here, $C_{nv}$ is the psychological cost derived from not acting upon betrayal and $D_{nv}$ is the utility derived from not voting (e.g. utility from not opposing a candidate). We assume that in subsequent elections, $D_{nv} > 0$ for a voter who previously voted in favor of an untrustworthy candidate. Thus for these voters, not voting in re-elections represents a form of implicit punishment.

We allow for the possibility that a partisan voter who did benefit from a winner's betrayal might still incur some psychological costs $C_{nv} \geq 0$ from not punishing that winner, but for this partisan voter $D_v < D_{nv}$.

The second social preference component captures voters' preferences for equality. Parameter $e \in [0; 1]$ captures how important it is for voters that wealth is distributed equally among voters. If equality in wealth distribution is important, then a voter who deems the probability of being pivotal $p$ to be high enough, receives additional utility from voting. This additional utility comes from reducing the inequality among partisan voters across periods 1 and 2, and is proportional to the absolute difference between the payoffs between the A-partisans and B-partisans, namely $\sum_{t=1}^{2} |\nu_{At} - \nu_{Bt}|$.

## 2.1 Indifferent voters

We can now concentrate on indifferent voters' behavior, who are our main object of inquiry.

While partisan voters always have an interest in seeing one candidate winning over the other[9] indifferent voters' monetary payoffs are always unaffected by electoral outcomes. This means that if $i = V_I$, then $\forall t$, $p \cdot \nu_{V_I t} = 0$.

An indifferent voter $V_I$ thus will vote if and only if:

$$\sum_{t=1}^{2} (g_{it} - c_{it}) + b \cdot \text{MAX}(D_v; D_{nv} - C_{nv}) + e \cdot \text{MAX}(p \cdot \sum_{t=1}^{2} |\nu_{At} - \nu_{Bt}|; 0) \geq 0 \qquad (2)$$

Equation (2) states that indifferent voters' choice to vote in the two elections depends only on their warm-glow preferences and social preferences. In particular, this equation suggests that in election 1 an indifferent voter's sole motivation to vote is warm-glow, since betrayal and wealth redistribution have not taken place yet. Moreover, if that voter chooses to vote, he faces no a-priori reason to favor one candidate over another.

---

[8] $D_v$ can be thought of as the benefit from "doing the right thing," or as the benefit from enforcing compliance of social norms.

[9] That is, if $i = V_A, V_B$ then, $\forall t$, $p \cdot \nu_{V_I t} > 0$.

Define the action set available in election 1 as $v_{it} = \{0; x; y\}$, where 0 indicates that the individual does not vote, and $x, y$ indicate that the individual votes for candidate $x$ or candidate $y$ respectively. This allows us to state our first proposition.

**Proposition 1.** *In the first election, an indifferent voter $V_I$ will vote if and only if $g_{V_I 1} \geq c_{V_I 1}$. If $g_{V_I 1} > c_{V_I 1}$, then $Pr(v_{V_I 1} = x) = Pr(v_{V_I 1} = y)$ regardless from what candidates promise, since voters have no preferences for either plan.*

In election 2, the choice to vote of an indifferent voter may depend on two additional factors: (1) whether the winner from election 1 maintained his promise; and (2) whether the incumbent's challenger, candidate $z$, promises a platform different from the one the incumbent chose to implement after winning election 1.

If indifferent voters care about politicians keeping their campaign promises, then a betrayal by the winner of the first election, in $t = 1$, may generate an additional incentive to vote in the second election, $t = 2$. This time however, all else equal, the recipient of their votes is not randomly determined as in election 1, because now altruistic punishment enters as one determinant of the vote choice. Because of altruistic punishment, the recipient of the ballots of these voters will be the new candidate who runs against the incumbent who broke his campaign promise.

Even when the incumbent from election 1 maintains his promise, an indifferent voter might still have an incentive to vote against this incumbent in election 2. This is due to preferences for equality.

Suppose in fact that the incumbent from election 1 promised and then implemented platform $A$, and he would also implement platform $A$ in $t = 2$, should he win in the second election. If the new challenger $z$ promises in $t = 2$ to implement platform $B$, then an indifferent voter who cares about equality of payoffs may vote for candidate $z$. The reason is that if platform $A$ is implemented in period $t = 1$, A-partisans receive a higher total payoff than B-partisans. By voting against the incumbent in $t = 2$, an indifferent voter can increase the probability that B-partisans will enjoy a higher payoff in $t = 2$, reducing thus the inter-temporal inequality of payoffs in the electorate. Put differently, an indifferent voter may vote to ensure that partisan citizens "take turns" in benefiting from politics.

More formally, define $W = \{x, y\}$ the winner of election 1 (either candidate $x$ or $y$), $p(W) = \{A, B\}$ the promise made by the winning candidate in election 1, $W_1 = \{A, B\}$ the effective decision made by the winning candidate, and $p(z) = \{A, B\}$ the promise made by challenger $z$ in election 2.[10]

Our second and fourth propositions describe our prediction regarding the behavior of indifferent voters in the second election, when the winner of the first election keeps his election campaign promise.

**Proposition 2.** *Suppose an indifferent voter does not care about equality, that is, $e = 0$. If $b = 0$, an indifferent voter $V_I$ will vote in election 2 if and only if $g_{V_I 2} > c_{V_I 2}$ and will*

---

[10]We assume that voters always trust candidate $z$'s promise.

*vote randomly, that is, $Pr(v_{V_I2} = W) = Pr(v_{V_I2} = z)$.*

Our third proposition presents the altruistic punishment hypothesis.

**Proposition 3.** *Suppose an indifferent voter does not care about equality, that is, $e = 0$. If $b = 1$ and $D_v > D_{nv} - C_{nv}$, then voter $V_I$ will punish the incumbent $W$, that is, $Pr(v_{V_I2} = W) = 0$ regardless of what candidate $z$ promises.*

The fourth proposition states that when indifferent voters care about equality of payoffs and the incumbents have kept their promise, the voting behavior of these indifferent voters depends on the platform choice of candidate z in the second election.

**Proposition 4.** *Suppose an indifferent voter does care about equality, that is, $e = 1$. If $b = 0$ and $W_1 \neq p(z)$, an indifferent voter $V_I$ will vote for candidate $z$. If $W_1 = p(z)$ and $g_{V_I2} > c_{V_I2}$, he will vote randomly, that is, $Pr(v_{V_I2} = W) = Pr(v_{V_I2} = z)$.*

The fifth hypothesis describes predictions when incumbents break their campaign promises and when voters value equality of payoffs among the electorate.

**Proposition 5.** *Suppose an indifferent voter does care about equality, that is, $e = 1$. If $b = 1$ and $W_1 = p(z)$, then if $D_v > D_{nv} - C_{nv}$, voter $V_I$ will punish the incumbent $W$, that is, $Pr(v_{V_I2} = W) = 0$. If $W_1 \neq p(z)$, then voter $V_I$ will still vote against the incumbent $W$ but the motivation is ambiguous, since it could be driven by either altruistic punishment, or inequality aversion, or both of these factors.*

## 3   Experimental Design

To test the hypothesis that voters use costly voting as a means of punishment, we design a voting environment that includes two key features. One feature is that some voters are monetarily indifferent between candidates running for electoral office. We refer to these voters as non-partisan or as indifferent voters. The other feature is that at least one candidate can break a campaign promise. These features allow us to identify if proposition 3 or proposition 5 or both propositions are consistent with our experimental data.

The first feature provides us with voters who are unaffected by a monetary interest in the election outcome. This implies that these voters are indifferent between election outcomes, and thus have no monetary reason for voting. Traditional rational choice models predict that these indifferent voters have no incentive to cast a ballot for or against either of the candidates in an election.

The second feature, that is the breaking of a campaign promise, implies the violation of a social norm by one of the candidates. For voters who are indifferent between the two candidates with respect to the monetary payoffs, this norm violation by one of the two candidates is the only distinguishing characteristic between the two candidates. Therefore,

we can attribute any voting decisions by indifferent voters that favors one over the other candidate to the norm violation. We integrate these two features in the following design.

In each session of ten participants, subjects are randomly assigned to the roles of voters or candidates. Of the ten participants, seven are voters and three are candidates. A session consists of two sequential elections. We randomly select two of the three candidates to participate in the first election. The third candidate is a challenger who runs in the second election against the winner of the first election, that is, the incumbent. It is common knowledge to all participants that the challenger in the second election, meaning the candidate who does not participate in the first election, will earn in that first election the same amount as the winning candidate of the first election. This aspect of our experiment assures that a desire to balance earnings between candidates does not influence voters' decisions, as might be the case if voters are inequity averse.

## 3.1 First Election

Voters and candidates know that the winning candidate of the first election will choose between two plans. Both plans divide a fixed amount of money (E$31) among voters. Table 1 shows that these plans are called "A1" and "B1." Both plans differ in how the E$31 is allocated to each voter. These payment plans represent the platforms or policy proposals of candidates. Candidates have perfect knowledge of how each plan allocates the E$31 between the seven voters. Voters have perfect knowledge of their own payment from each plan, and the payment to other voters from each plan. Voters do not know about candidate payoffs.

The payoff matrix in Table 1 shows that each of the two payment plans or policy proposals allocates the same earnings to five out of the seven voters. These five voters earn E$ 4 no matter which plan is implemented by the winner of the election. Because of this payoff structure, we refer to these voters as non-partisan or indifferent voters. Further, voting is costly. Casting a vote costs E$0.50. Thus, these five voters have no monetary incentive to cast a vote because voting is costly.

In contrast to these five indifferent voters, two voters have preferences for one or the other candidate. We refer to these two voters as "partisans." One of these partisans prefers payment plan A1, because that payment plan gives that voter a higher payoff than payment plan B1. Similarly, the other partisan prefers payment plan B1, because that payment plan gives that voter a higher payoff than payment plan A1. We chose the payoffs shown in Table 1 so, that the preferences of these two voters are opposite, but symmetrical regarding what plan they prefer. The symmetry in payoffs implies that each partisan has the same high payoff when the candidate who wins the election chooses one of the plans. This design feature implies that an indifferent voter with altruistic preferences has no incentive to vote for one or the other candidate in election 1. This is because social welfare, measured as total payoffs, is identical regardless of which candidate wins the election.

9

Table 1: Voters' payoffs in Election 1

|  | Payment Plan | |
| --- | --- | --- |
|  | A1 | B1 |
| Voter 1 | 4 | 4 |
| Voter 2 | 4 | 4 |
| Voter 3 | 4 | 4 |
| Voter 4 | 4 | 4 |
| Voter 5 | 4 | 4 |
| Voter 6 | 9 | 2 |
| Voter 7 | 2 | 9 |
| Total Payment | 31 | 31 |

Prior the first election, both candidates send a non-binding (cheap-talk), closed-form message to all voters. In this message they announce which plan they would implement if elected. Both voters and candidates know that this message is not binding. Once the two candidates have sent their messages, each voter chooses between not voting and casting a vote in favor of one of the two candidates.

The winning candidate is the candidate who receives the largest number of votes. The winner then chooses between payment plan A1 and B1. In our baseline design, we inform all voters of the winning candidate's choice, and we inform voters about the winning candidate's payoff incentives regarding his choice which of the two payment plans to implement. In the baseline design, if the winning candidate breaks his promise he is paid $20, otherwise he is paid $4. This feature of our experiment provides winners in the first election with a strong incentive to break their promise.

Table 2: Winning candidate's payoffs in Election 1

|  | Promise A1 | Promise B1 |
| --- | --- | --- |
| Implement A1 | 4 | 20 |
| Implement B1 | 20 | 4 |

In addition to the earnings from the implemented payment plan, all voters receive an additional 2E$ independent of the chosen payment plan. All participants know about this additional payment at the beginning of the experiment, and know that this plan is paid regardless of whether they choose to vote or abstain.

## 3.2 Second Election

In the second election, the winner of the first election, that is, the incumbent, and the challenger, who is the third candidate who did not participate in the first election, compete for voters' ballots. Table 3 shows the two payment plans available in this second election. These two payment plans have the identical payoffs as the payment plans available in the first election.

In this second election, the incumbent does not send a new promise. It is common knowledge to voters and to candidates that in the event that he or she wins the second election, the incumbent is bound to implement the same payment plan that he or she chose after the first election. For example, if the winner of the first election chose payment plan A1 (B1) after the first election, then the incumbent is required to choose A2 (B2) if he wins the second election.

The challenger in the second election sends a non-binding message just as the two candidates did in the first election. The two payment plans in the second election, "A2" and "B2," are identical to the two payment plans "A1" and "B1" in the first election.

Table 3: Voters' payoffs in Election 2

|  | Payment Plan | |
| --- | --- | --- |
|  | A2 | B2 |
| Voter 1 | 4 | 4 |
| Voter 2 | 4 | 4 |
| Voter 3 | 4 | 4 |
| Voter 4 | 4 | 4 |
| Voter 5 | 4 | 4 |
| Voter 6 | 9 | 2 |
| Voter 7 | 2 | 9 |
| Total Payment | 31 | 31 |

Prior to the voting phase of the second election, we inform all subjects that the challenger was paid the same amount as the incumbent in the first election. This ensures that inequity concerns between the initial payments of the incumbent and challenger are not influencing voting behavior in the second election.

In this design, indifferent voters have no monetary incentive to vote against the incumbent, even if the incumbent reneged on his first election promise. Further, voters know that our experiment requires candidates to send a message, so sending a message was not their choice. Additionally, voters know that we give first election winners an incentive to betray. They receive $4 if they do not betray and $20 if they do betray. Further, voters know that betrayal increases the aggregate payoff by 16E$, thus that betrayal increases

social welfare.

After the first and after the second election we incentivized subjects to estimate the correct number of voters who voted for the winning candidate. We asked them to do so after announcing the winner, but before revealing winning candidate's decision. We rewarded correct estimates with 5 E$. Since voters were asked to estimate the number of voters before winner's decision, and thus prior to any potential betrayal, subjects have an incentive to truthfully reveal their priors in the first election.

## 3.3   Treatments

Our experiment consists of three treatments. First, in the BASELINE treatment candidates have an incentive to break their promise to voters. We described this treatment in the previous sub-section of this paper. Second, in the NOLIES treatment, candidates have no incentive to break their campaign promises. This treatment provides us with an environment that allows us to explore how indifferent voters make decisions when not faced with broken promises. Third, in the INVEST treatment, candidates have an incentive to break their campaign promises and, differently from the baseline, a subset of voters have opportunity to invest in candidates' promises, prior to learning whether the promise was broken. As we describe below, this treatment brings our baseline a step closer to voting in the field[11], and tests whether we do still find consistent patterns of altruistic punishment.

### 3.3.1   BASELINE Treatment

In BASELINE, the winning candidate has a monetary incentive to implement the plan that he did not signal to voters prior to the election that he or she would choose if elected. Table 2 shows that the winning candidate faces this incentive because he earns E$20 for implementing the payment plan opposite to the one he had promised, and earns E$ 4 for implementing the payment plan he had promised during the campaign. We do not inform voters about this payoff structure until after the winner choses a payment plan after having obtained the majority of votes in the first election.

Since indifferent voters receive the same earnings in plans A1 and B1 and the same earnings in plans A2 and B2, they have no monetary incentive to cast a ballot in either election. Further, any implicit utility derived from the act of voting itself, due to habits or civic duty, for example, will lead indifferent voters to vote indiscriminately between candidates in the first and in the second elections.

The voting behavior of indifferent voters can be interpreted as punishment voting behavior when we observe one of three within-subject voting patterns. First, financially indifferent voters might systematically switch from voting for the incumbent in the first election to voting for the challenger in the second election. Second, financially indifferent

---

[11]Even politically indifferent individuals in fact may still take decisions based on the expectation that political promises will be kept.

voters might systematically switch from voting for the incumbent in the first election towards abstaining in the second election. In this voting pattern, we infer that abstaining voters incur a cost because in the first election these voters had expressed a preference for the act of voting. This is because they had cast a ballot in the first election even though they were indifferent between the two possible outcomes of the election.[12] Third, financially indifferent voters might systematically switch from abstaining in the first election to voting for the challenger in the second election. This type of punishment voting by these voters is costly. And since this voting is associated with a norm violation that harmed another individual and not the financially indifferent voters themselves, we consider this third type voting pattern to be altruistic punishment. This type of punishment is similar to third party-punishment after subjects have played a dictator game (see i.e. Fehr and Fischbacher 2004), a trust game (Charness, Cobo-Reyes and Jimenez 2008), or public goods game (e.g. Carpenter and Matthews 2004).

To identify whether indifferent voters use their votes as punishment, our design excludes the possibility that voters cast their ballot against the incumbent due to inequity aversion or inequity concerns with respect to candidate payments. We remove this potential inequity concern by paying the challenger the same amount as the winning candidate in the first election( either E\$ 4 or E\$ 20, depending on winners decision). Moreover, we make this information common knowledge.

While our design excludes the possibility that inequality of payoffs among candidates influences voting decisions, there is the potential for an inequality of payoffs between partisan voters. For example, if the winning candidate chooses payment plan A1 in the first election, one partisan voter receives the higher payoff E\$9 and the opposing partisan receives the lower payoff E\$2. Should the incumbent win again in the second election, the incumbent is required to implement the same plan as in the first election, leading to an increase in the inequity of payoffs between the two partisan voters. An inequality averse indifferent voter might cast a ballot against the incumbent, not as punishment but in hopes that the challenger chooses the alternative plan, which would give the previously low earning partisan the greater payment and the previously high earning partisan the lower payment. This would result in equity of payments between the two partisans over the two elections. We design our second treatment, the NOLIES treatment, to disentangle the motivation to cast a ballot to punish from the motivation to reduce inequity of payoffs amoung partisans.

### 3.3.2  NOLIES Treatment

In the NOLIES treatment, we remove the incentive for candidates to break their promise, thereby effectively eliminating the punishment motive for indifferent voters. In the NO-LIES treatment the only change relative to the BASELINE treatment is that candidates

---

[12]As our theoretical framework indicates, this action implies that $g_{it} > c_{it}$.

face no monetary incentives to act contrary to their promise. No matter what the winning candidate promises and implements, that candidate always earns E\$20. Because the winning candidate has no incentive to break his promise, he has little reason to choose the other plan. In this design, we have no altruistic punishment motive with respect to voting against the incumbent in the second election and thus we predict no voting due to altruistic punishment.

Voters in the NOLIES treatment face the same payment plans as in the BASELINE treatment. As in the BASELINE treatment, if the incumbent wins the second election, he is required to implement the same payment plan in the second election as the plan after the first election. Thus, the same inequality concern that would drive an inequity averse indifferent voter to vote against the incumbent in BASELINE, also would motivate an inequity averse indifferent voter to cast a ballot against the incumbent in NOLIES. Therefore, if inequity aversion is the motivating influence for indifferent voters to vote against the incumbent in BASELINE, then we will observe a similar voting pattern in the NOLIES treatment when the incumbent no longer breaks his promises. If, on the other hand, indifferent voters are less likely to vote against the incumbent in the second election of the NOLIES treatment, then inequity aversion cannot explain the observed voting pattern in the BASELINE treatment.

### 3.3.3 INVESTMENT Treatment

In the field, voters do not know immediately after the election whether politicians will deliver on or renege on a campaign promise. In this time prior to policy implementation, voters who had been indifferent between payment plans may make financial decisions based upon campaign promises. If an elected candidate then breaks that campaign promise, individuals who made decisions based on the assumption that a promise will not be broken might suffer losses.

Our third treatment, INVEST, tests the hypothesis that indifferent voters who make an investment based on a campaign promise are more likely to engage in costly punishment when that promise is broken. In this third treatment, the indifferent voters are no longer third-party punishers but first-party punishers, because they are worse off when incumbents renege on their campaign promises.

The INVEST treatment builds on the BASELINE treatment. Relative to the BASELINE treatment, the INVEST treatment provides a subset of voters with an option to invest the additional 2E\$ that each subject receives after each election.

In the INVEST treatment, five out of the seven voters receive the option to invest 2E\$. If they choose not to invest, they will keep the 2E\$ regardless of the winning candidate's choice. If they choose to invest, their investment return will depend on the winning candidate's decision. If the winning candidate breaks his campaign promise then the investment return is zero and the indifferent votes loses 2E\$. If the winning candidate keeps his campaign promise, then the investment return is 100 percent and the investor receives 4E\$ from

14

their 2E$ investment. Voters make their investment decisions after the announcement of the winning candidate, but prior to the announcement of the winning candidate's payment plan decision.

# 4    Results

We conducted our experiment at George Mason University with 260 undergraduate participants. Upon arrival at the lab, we randomly assigned participants to the role they will assume in the experiment, that is, either to the role of a voter or that of a candidate. We ran 26 sessions, with ten participants in each session. The 26 sessions include eleven sessions (110 subjects) in the BASELINE treatment, five sessions (50 subjects) in the NO-LIES treatment, and ten sessions (100 subjects) in the INVEST treatment. This provides us with voting decisions of 130 indifferent voters.

## 4.1    Summary

Since our three treatments are nearly identical with the exception of either whether the candidate has an incentive to betray and whether a subset of voters can invest based on the campaign promise, we start exploring results by pooling our data across the three treatments. In this initial analysis of the data, we investigate how behavior differs when the candidates break their promises or do not break their promises.

**Result 1.** *Most indifferent voters do not vote.*

As predicted, indifferent voters do not tend to vote in the first election. Pooling findings from when the winner broke and kept his campaign promise, we find that 69.2% of indifferent voters, or 90 out of 130 voters, abstain in the first election. This suggests that 40 of the 130 indifferent voters (30.8%) have an intrinsic motivation for voting and that they value voting at or above the monetary cost of voting, which we had set to 0.50 E$. Further, we find that 61.5% of indifferent voters (80 out of 130) do not vote in both the first and second election.

Table 4 shows the raw and percentage numbers regarding the behavior of indifferent voters. The table breaks down this behavior by whether incumbents broke their campaign promise (top panel of Table 4) or kept their campaign promise (bottom panel of Table 4). We observe, across our three treatments, that twenty incumbents break their campaign promise and six incumbents keep their promise. Further this table breaks down the behavior of indifferent voters by whether it pertains to the first election (first column of Table 4) or to the second election (second column of Table 4).

The top panel of Table 4 shows that 100 indifferent voters were in a treatment where the incumbents did not keep their campaign promises. The bottom panel of Table 4 shows that 30 indifferent voters were in a treatment where incumbents kept their campaign promise.

Table 4: Summary Statistics for Indifferent Voters in All Elections

| First Election | | | Second Election -Winner Broke Promise | | |
|---|---|---|---|---|---|
| | n. | % | | n. | % |
| **Vote** | 34 | (34%) | **Vote for Incumbent** | 12 | (12%) |
| | | | **Vote for Challenger** | 25 | (25%) |
| **Abstain** | 66 | (66%) | **Abstain** | 63 | (63%) |
| **All Indifferent Voters** | 100 | (100%) | **All Indifferent Voters** | 100 | (100%) |

| First Election | | | Second Election -Winner Kept Promise | | |
|---|---|---|---|---|---|
| **Vote** | 6 | (20%) | **Vote for Incumbent** | 3 | (10%) |
| | | | **Vote for Challenger** | 3 | (10%) |
| **Abstain** | 24 | (80%) | **Abstain** | 24 | (80%) |
| **All Indifferent Voters** | 30 | (100%) | **All Indifferent Voters** | 30 | (100%) |

**Result 2.** *A significant portion of indifferent voters are willing to vote as altruistic punishment.*

The first column of Table 4 shows, as predicted, that roughly the same proportion of indifferent voters abstain regardless of whether the winning candidate eventually breaks their promise. That is, the 66 percent abstaining prior to a candidate following through on a promise and the 80 percent abstention rate prior to a candidate breaking a promise are not significantly different (Two-tailed Mann Whitney test, p>0.10).

However, the voting behavior differs during the second election based upon whether winners broke or kept their promises. In the second election, an equal proportion votes for the challenger and incumbent when incumbents kept their promise (10% of the indifferent voter population for each candidate). However, when incumbents broke their promise, more than two thirds of those voting favor the challenger in the second election. Put differently, 25 percent of all indifferent voters favor the challenger and 12 percent of all indifferent voters favor the incumbent.

The findings in this Table 4 allow for an initial test our hypothesis that indifferent voters are willing to engage in costly voting as a punishment device. If indifferent voters have no punishment preferences, they will spread their votes equally between the challenger and the incumbent. However, the results in top panel of Table 4 for the second election show that this is not the case. Since more than two-thirds of the voting indifferent voters are choosing to vote for the challenger after the incumbent breaks his campaign promise,

we find support for the hypothesis that voters using voting as a punishment device. This effect is statistically significant (two-tailed t-test, p=0.03).

As discussed above, favoring the challenger in the second election is also consistent with inequity averse preferences. However, if inequity averse preferences motivate the observed behavior we would also observe a shift towards voting for the challenger in the second election when the incumbent does not break his or her promise. Instead, the results for the second election in the bottom panel of Table 4 shows that exactly half of the voting indifferent voters cast a ballot for the incumbent and half for the challenger when incumbents do not break their promises. This further supports our first evidence of the existence of voting as altruistic punishment, as stated in Result 2.

## 4.2  Identifying Punishment and Inequity Averse Voting Preferences

The voting patterns of indifferent voters provide a more detailed look at altruistic punishment voting. Table 5 breaks the behavior reported in Table 4 into more detail. The first column of Table 5 shows the voting behavior of indifferent voters in the first election, and thus replicates the first column of Table 4. As in that table, the top panel of Table 5 shows the results for the second election when the winners of the first election did break their campaign promises, and the bottom panel shows the corresponding information when winners of the first election did not break their campaign promises.

Voters who always abstain in elections, or voters who always vote for the challenger, or always vote for the winner of the first election do not help us to make inferences regarding altruistic punishment. Columns 2 and 3 of Table 5 show how many voters and the percentage of these votes that we have in each of these categories. Column 2 shows that as a percentage of those who abstained in the first election, 86 percent of our indifferent voters abstain also in the second elections when the incumbent broke their promise. Ninety-six percent of those who abstained in the first election also abstained in the second election when the incumbent did not break their promise.

Table 5, Column 3 shows that in the broken promise treatment 47 percent of voters who voted for a candidate in the first election, cast a ballot for the incumbent in both elections, or for the losing candidate in the first election and the challenger in the second election. We find that a treatment with no broken promises, this percentage increases to 67 percent.

The bottom panel of Table 5 shows that among the 30 indifferent voters who did not observe a winning candidate break their promise, 23 subjects or 76.7% of these voters abstained (Table 5, column 2) in both elections. Four voters, or 13.3%, voted either for the incumbent in both elections or for his opposing candidate in both elections (Table 5, column 3).

To better understand any punishment motivations for voting, columns 4 to 9 of Table 5 show the voting behavior of indifferent voters in the second election who changed their voting behavior relative to that in the first election, broken into descriptive categories. These

columns provided data that describe punishment preferences in voting. These columns provide information about the aforementioned three patterns of punishment and inequity averse voting (from section 3.3.1) and help us to distinguish between the punishment and inequity voting motivations.

Columns 4, 5, and 6 indicate several voting patterns that are consistent with indifferent voters switching away from the incumbent towards the challenger. Column 4 shows the fraction and the percentage of voters who switched from voting for the incumbent in the first election to voting for the challenger in the second election. Column 5 shows the fraction and percentage of voters who switched from voting for the incumbent in the first election to abstaining in the second election. This behavior reduces the number of votes for the incumbent. It also demonstrates a costly change in voting since their first election vote demonstrated a revealed intrinsic value of voting greater the voting cost, which they now forgo in the second election. Finally, column 6 shows the fraction and percentage of voters who initially abstained from voting in the first election and who then switch to voting for the challenger in the second election. This voting behavior demonstrates a revealed lack of an intrinsic value of voting (due to their abstention in the first election) and a costly decision to cast a ballot in the second election. In other words, a voter appearing in any of these three columns in the top panel of the table is consistent with the voter acting from either or both inequity aversion and/or altruistic punishment preference motivations, whereas a voter appearing in the bottom panel of the table is consistent with acting only from an inequity aversion motivation.

Columns 7, 8, and 9 show the voting patterns that are consistent with switching away from the challenger, and are analogous to columns 4,5, and 6. Column 7 shows those voters who switch from voting for the challenger to voting for the incumbent. Column 8 shows voters who abstained after voting for the candidate who lost the first election. Column 9 shows those that abstained in the first election and voted for the incumbent in the second election. These columns show voting patterns that are not consistent with inequity averse preferences or altruistic punishment preferences, since relative shifts towards the incumbent would increase the chance of greater inequality.

Since our focus is on altruistic punishment voting and possible inequity aversion motivations for voting we henceforth concentrate our attention on voting patterns reflected in columns 4, 5, and 6. As previously mentioned, the inequity aversion motivation for voting in one of these three ways is present in both the top panel (second elections after a broken promise) and in the bottom panel (second elections after a kept promise).[13]

Punishment preference motives for voting patterns 4, 5, and 6 would only be present in the top panel, when the incumbent broke their promise previously. Therefore, the

_____

[13]Random voting decisions, partially indicated by voting patterns in columns 7, 8, and 9, would also be thought to be constant between top and bottom panels. We find that there is no significant difference in the voting patterns in columns 7, 8 and 9 following broken or kept promises (Two-tailed Mann Whitney, p=0.34). The 3% of indifferent voters voting in one of the patterns from columns 7 to 9 seen in the top panel is marginally different from zero (two-tailed t-test, p=0.08).

Table 5: Indifferent voters' behavior in second elections conditional on first elections' behavior (all treatments)

| First Election | | | | Second Election – When Winner Broke Promise | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| | | Abstain in Both Elections | Vote for Same Candidate (Winner/Incumbent or Loser/Challenger) | Switch Vote To Challenger | Abstained After Voted for Winner | Switched from Abstaining to Voting for the Challenger | Switch Vote To Incumbent | Abstained After Voted for Losing Candidate | Switched from Abstaining to Voting for the Incumbent |
| Vote | 34 34.0% | - | 16/34 47.1% | 12/34 35.3% | 4/34 11.0% | - | - | 2/34 5.9% | - |
| Abstain | 66 66.0% | 57/66 86.4% | - | - | - | 8/66 12.1% | - | - | 1/66 1.5% |
| All Indifferent Voters | 100 100.0% | 57/100 57.0% | 16/100 16.0% | 12/100 12.0% | 4/100 4.0% | 8/100 8.0% | 0/100 0.0% | 2/100 2.0% | 1/100 1.0% |

| First Election | | | | Second Election – When Winner Did Not Break Promise | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Abstain in Both Elections | Vote for Same Candidate (Winner/Incumbent or Loser/Challenger) | Switch Vote To Challenger | Abstained After Voted for Winner | Switched from Abstaining to Voting for the Challenger | Switch Vote To Incumbent | Abstained After Voted for Losing Candidate | Switched from Abstaining to Voting for the Incumbent |
| Vote | 6 20.0% | - | 4/6 66.7% | 1/6 16.7% | 1/6 16.7% | - | - | - | - |
| Abstain | 24 80.0% | 23/24 95.8% | - | - | - | 1/24 4.2% | - | - | - |
| All Indifferent Voters | 30 100.0% | 23/30 76.7% | 4/30 13.3% | 1/30 3.3% | 1/30 3.3% | 1/30 3.3% | 0/30 0.0% | 0/30 0.0% | 0/30 0.0% |

*Note: Observations in columns (4), (5) and (6) are consistent with both altruistic punishment and inequity aversion.*
*Observations from columns (7), (8), (9) are consistent with random voting but not with altruistic punishment and inequity aversion.*

difference in the percentage of column 4, 5, or 6 pattern voters when a promise was broken (12%+4%+8% =24%, top panel) from the percentage when a promise was unbroken (3.3%+3.3%+3.3% = 10%, bottom panel) gives an additional estimate of 14% (24% -10%) for the fraction of indifferent voters with altruistic punishment motives for voting.[14]

This leads us to result 3.

**Result 3.** *Our data estimates that at least 14% of indifferent voters have voting patterns that are consistent with altruistic punishment voting.*

While inequity aversion is the only factor that might lead an individual to cast a ballot in the case where no promise was broken, an inequity-averse individual may also have punishment preferences. Therefore, the 14% of voters who are willing to incur a cost to express punishment preferences are those voters with punishment preferences and no inequity aversion preferences. Thus, this fraction of voters describes only a lower bound of the fraction of individuals with punishment preferences. Including voters with both preferences will increase the fraction of voters who are willing to incur a cost to cast a ballot against the promise-breaking incumbent.

As a robustness test to ensure that the behavior in Table 5, and in particular that the behavior shown in columns 4, 5, and 6, is not just reflecting random voting patterns, we employed a Monte Carlo simulation of the behavior in the second election. We first fixed the first round behavior for a group of 100 simulated indifferent voters to be the same as the decisions of the 100 indifferent voters in the experiment, that appear in the top panel of the table. That is we assure that we have the same frequency of abstaining or voting for a winner or loser. We then simulated the 100 simulated indifferent voters' decisions to abstain, vote for incumbent, or vote for the challenger in the second election. To generate a Monte Carlo sample, we randomly permuted the choice behavior in the second election across 100 simulated indifferent voters. We simulated the second election voting behavior for 100,000 such 100-person groups.

The summary statistics computed using these Monte Carlo samples reflect the null distribution, that we then compare to our own sample behavior, to evaluate the probability our data reflects random voting. For instance, if a simulated voter abstained in the first election, that voter would have three possible ways to vote in the second election: abstain a second time (appear in column two); switch to voting for the challenger (appear in column six); or switch to voting for the incumbent (appear in column nine).

The Monte Carlo simulation provides evidence that our observed data is statistically significantly different from random voting behavior (two-tailed, p<0.02) in all cases but two. That is, there are two situations where the Monte Carlo simulation suggests we cannot reject the possibility that those voters were voting randomly. One situation is that in column 7 (voters switching to voting for the incumbent, p=0.79) and the second situation is that in column 8 (voters abstaining after voting for the original losing candidate, p=0.13). Since a theory of altruistic punishment voting does not predict systematic voting in these cases anyway, the Monte Carlo evidence supports the conclusion that our data reflect systematic voting patterns as discussed in our results above.[15]

---

[14]We can demonstrate the statistical significance of this 14% in a number of ways. First we observe that the proportion of people voting in categories (4)-(6) when the incumbent betrays is statistically significantly greater than the 10% observed when the incumbent does not betray (two-tailed t-test using hypothesized mean of 10%, p<0.01; one-tailed Mann-Whitney p<0.05). Further a Chi-Squared test of the distribution of voting when previously betrayed is statistically significantly different than the distribution of voting when not betrayed, regardless of whether or not we pool abstaining twice and voting for the same candidate twice (Chi-Squared, p<0.01 under either distribution).

[15]For columns 2 through 9 the resulting p-values derived from the 100,000 Monte Carlo simulations of the top panel of Table 5 are as follows: 0.000 (Column 2), 0.000 (Column 3), 0.016 (Column 4), 0.000 (Column 5), 0.000 (Column 6), 0.792 (Column 7), 0.127 (Column 8), 0.019 (Column 9)

## 4.3  Investment Opportunity

In this section, we no longer perform a non-parametric analysis and now employ regression analysis to explore the existence of punishment voting. The approach in this section also provides us with means to analyze the influence of the betrayal's salience on punishment decisions.

Table 6:  Probit Analysis of Punishment Consistent with Voting Patterns (types: (4), (5), (6))

|  | Excluding INVEST Treatment | All Treatments Included |
|---|---|---|
| Candidate Broke Promise | 0.17** | 0.16** |
|  | (0.07) | (0.06) |
| Beliefs | 0.07** | 0.07*** |
|  | (0.02) | (0.02) |
| Investment Treatment | – | -0.07 |
|  |  | (0.07) |
| Invested | – | 0.16 |
|  |  | (0.15) |
| Observations | 80 | 130 |
| Prob>Chi$^2$ | 0.015 | 0.001 |
| Log pseudolikelihood | -32.759 | -56.480 |

*Note: ** implies significance at a 5% level and *** at 1% level.*
*"Beliefs" represents subjects saliently rewarded belief*
*immediately after the first election of how many people*
*voted for the winning candidate.*
*Coefficients are marginal probabilities.*

Table 6 presents a probit analysis predicting the probability that an indifferent voter will vote in a manner consistent with altruistic punishment or inequity aversion, in other words the probability that an indifferent voter will vote in patterns 4, 5, or 6 as seen in Table 5. Thus, in our Probit analysis, the dependent variable equals one if the voting behavior is consistent with punishment preferences (types (4),(5),(6)), and zero otherwise. If such voting is reflecting inequity aversion alone then the "Candidate Broke Promise" variable (an indicator variable equal to one if the candidate broke their promise) should not provide any predictive power. However if broken promises result in significantly increase in such voting patterns, then this is evidence that is consistent with of altruistic punishment voting. The Probit analysis also provides an opportunity to incorporate the beliefs of voters on

how many people voted for the incumbent in the first election, which could not be easily incorporated into our previous analyses. Voters might shift away from voting, to abstain for instance, if they believe that a large fraction of voters already voting in the way they prefer, especially since voting is costly.

Table 6 also provides us the first opportunity to explicitly look at the INVEST treatment as being potentially different from the BASELINE. The first column excludes the INVEST treatment in the analysis, and the second column includes the INVEST treatment as indicated. Remember, indifferent voters in the INVEST treatment, as with all of our treatments, receive the same payoff under all payment plans, both in the first and the second election. After the first election however, they now have a chance to invest the E\$ 2 in either campaign promise A1 or campaign promise B1.

Indifferent voters who take advantage of this investment opportunity suffer a monetary loss when a promise is broken, while they receive a 100 percent return when the campaign promise is implemented. While the loss of the E\$ 2 investment may give indifferent voters who invest an additional incentives to exercise punishment, their payoffs in the second election are unaffected by the election's outcome.

The probit analysis in Table 6 provides additional evidence for altruistic punishment The point estimates on whether the winning candidate in the first election broke his campaign promise are statistically significant in both models ($p< 0.05$). If indifferent voters were motivated solely by inequity aversion, this coefficient would be no different than zero.

Table 6 shows that there is not any statistically significant increase in the prevalence of punishment consistent voting behavior if an indifferent voter invested in a campaign promise or was in an environment where they knew other voters may have invested in a campaign promise($p >0.1$). This finding shows two things about the punishment motive for voting. First, it shows that this punishment motivation is unaffected by the degree of harm done to the voting population. As we mentioned, in the investment treatment, investors were harmed above and beyond any harm from the betrayal in the baseline treatment. However, the probit results show no effect of this extra harm on propensity to punish. Second, the lack of a statistically significant effect of one's own investment in the probit analysis suggests that the propensity to vote to punish is not affected by having personally been harmed by the betrayal. Taken together with the statistical significance of the broken promise variable, these two observations about investment support the hypothesis that punishment voters are voting as altruistic punishment and not out of a feeling of personal harm or vengeance.

Further, the results in the table show the significant influence of beliefs. Indifferent voters are significantly more likely to vote against that incumbent in the second election ($p<0.05$ in both models) when they expect that more people voted for the incumbent in the first election.[16]

---

[16]This statistically significant effect of beliefs is consistent with both inequity averse and altruistic punishment preferences.

# 5 Conclusion

In our experiments, a subset of the electorate is monetarily unaffected by proposed political platforms and associated promises, and candidates have the option to betray their electoral promises. If indifferent voters are solely motivated by their monetary payoff, a rational choice model predicts that these indifferent voters will not cast a ballot. And if these indifferent voters cast a vote, a rational choice model predicts that the distribution of these votes is unaffected by a candidate's betrayal. Our results are not consistent with these predictions. We document that at least 14% of indifferent voters utilize voting as a means of altruistic punishment. Our results provide the first evidence of the existence of altruistic punishment voting as well as the first estimate for the lower bound of voters who cast a costly vote, which is due to the desire to punish the incumbent for breaking his campaign promise. We show this effect is strongly significant even in an environment where breaking promises expands overall social welfare.

Our results have several implications for environments that use voting to elect candidates. First, our results suggest that citizens who are indifferent between political outcomes may have a significant role in swinging elections in disfavor of untrustworthy politicians. This suggests the existence of a mechanism by which an informed electorate in a democracy can prevent at least some corruption and dishonesty in the ruling body. Future studies may find it valuable to explore how the degree of information available to voters affects voting turnout. In our experiment in fact, subjects learn with certainty whether a candidate 'intentionally broke a promise. Further work may analyze the relative impact of altruistic punishment characterized by uncertainty about the intentionality of promise breaking.

Second, with higher voting frequencies, voters have more opportunity to punish incumbents who break their promises. Our results suggest that voters take advantage of this opportunity. Punishment voting in turn might give politicians who face more frequent elections an incentive to keep their campaign promises.

In this paper, we studied voting behavior in an environment with a single campaign promise. Clearly, candidates in elections make several promises, many of them being vague, some specific. Future research might include an examination of the effect of bundling vague campaign promises with specific campaign promises, or the bundling of several specific campaign promises, and study how voters react when one promise is broken but another is not. It may also be fruitful to explore how ?price sensitive? altruistic punishment motivations are to the costs associated with the voting environment. Our design may serve well as a base design model for testing either of these questions.

# References

[1] Alan I Abramowitz, David J Lanoue, and Subha Ramesh. Economic conditions, causal attributions, and political evaluations in the 1984 presidential election. *The Journal of Politics*, 50(04):848–863, 1988.

[2] Jason A Aimone and Daniel Houser. Harnessing the benefits of betrayal aversion. *Journal of Economic Behavior & Organization*, 89:1–8, 2013.

[3] André Blais. *To vote or not to vote?: The merits and limits of rational choice theory.* University of Pittsburgh Press, 2000.

[4] Iris Bohnet, Fiona Greig, Benedikt Herrmann, and Richard Zeckhauser. Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *The American Economic Review*, pages 294–310, 2008.

[5] Robert Boyd, Herbert Gintis, Samuel Bowles, and Peter J Richerson. The evolution of altruistic punishment. *Proceedings of the National Academy of Sciences*, 100(6):3531–3535, 2003.

[6] Geoffrey Brennan and James Buchanan. Voter choice: Evaluating political alternatives. *American Behavioral Scientist*, 1984.

[7] Geoffrey Brennan and Alan Hamlin. Expressive voting and electoral equilibrium. *Public Choice*, 95(1-2):149–175, 1998.

[8] Geoffrey Brennan and Loren E Lomasky. *Democracy and decision: The pure theory of electoral preference.* Cambridge University Press, 1997.

[9] Stephen Coate, Michael Conlin, and Andrea Moro. The performance of pivotal-voter models in small-scale elections: Evidence from texas liquor referenda. *Journal of Public Economics*, 92(3):582–596, 2008.

[10] Stefano DellaVigna, John A List, Ulrike Malmendier, and Gautam Rao. Voting to tell others. Technical report, National Bureau of Economic Research, 2014.

[11] Ingolf Dittmann, Dorothea Kubler, Ernst Maug, and Lydia Mechtenberg. Why votes have a value. 2007.

[12] Anthony Downs. An economic theory of democracy. 1957.

[13] Timothy Feddersen and Alvaro Sandroni. A theory of participation in elections. *The American Economic Review*, pages 1271–1282, 2006.

[14] Timothy J Feddersen. Rational choice theory and the paradox of not voting. *Journal of Economic Perspectives*, pages 99–112, 2004.

[15] Ernst Fehr and Urs Fischbacher. The nature of human altruism. *Nature*, 425(6960):785–791, 2003.

[16] Ernst Fehr and Urs Fischbacher. Social norms and human cooperation. *Trends in cognitive sciences*, 8(4):185–190, 2004.

[17] Ernst Fehr and Urs Fischbacher. Third-party punishment and social norms. *Evolution and human behavior*, 25(2):63–87, 2004.

[18] Ernst Fehr and Simon Gächter. Altruistic punishment in humans. *Nature*, 415(6868):137–140, 2002.

[19] Ernst Fehr and Simon Gächter. Human behaviour: Egalitarian motive and altruistic punishment (reply). *Nature*, 433(7021):E1–E2, 2005.

[20] Morris P Fiorina. The voting decision: instrumental and expressive aspects. *The Journal of Politics*, 38(02):390–413, 1976.

[21] Morris P Fiorina. Retrospective voting in american national elections. 1981.

[22] Morris P Fiorina. Information and rationality in elections. *Information and democratic processes*, pages 329–342, 1990.

[23] James H Fowler. Altruistic punishment and the origin of cooperation. *Proceedings of the National Academy of Sciences of the United States of America*, 102(19):7047–7049, 2005.

[24] James H Fowler. Habitual voting and behavioral turnout. *Journal of Politics*, 68(2):335–344, 2006.

[25] James H Fowler, Tim Johnson, and Oleg Smirnov. Egalitarian motive and altruistic punishment. *Nature*, 433(10.1038), 2005.

[26] Alan S Gerber, Donald P Green, and Christopher W Larimer. Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, 102(01):33–48, 2008.

[27] Benny Geys. rationaltheories of voter turnout: a review. *Political Studies Review*, 4(1):16–35, 2006.

[28] John C Harsanyi. Morality and the theory of rational behavior. *Social Research*, pages 623–656, 1977.

[29] Arye L Hillman. Expressive behavior in economics and politics. *European Journal of Political Economy*, 26(4):403–418, 2010.

[30] Tim Johnson, Christopher T Dawes, James H Fowler, Richard McElreath, and Oleg Smirnov. The role of egalitarian motives in altruistic punishment. *Economics Letters*, 102(3):192–194, 2009.

[31] Erik O Kimbrough, Vernon L Smith, and Bart J Wilson. Historical property rights, sociality, and the emergence of impersonal exchange in long-distance trade. *The American Economic Review*, pages 1009–1039, 2008.

[32] Donald R Kinder and D Roderick Kiewiet. Economic discontent and political behavior: The role of personal grievances and collective economic judgments in congressional voting. *American Journal of Political Science*, pages 495–527, 1979.

[33] John O Ledyard. The pure theory of large two-candidate elections. *Public Choice*, 44(1):7–41, 1984.

[34] Assar Lindbeck, Sten Nyberg, and Jörgen W Weibull. Social norms and economic incentives in the welfare state. *The Quarterly Journal of Economics*, 114(1):1–35, 1999.

[35] John G Matsusaka. Explaining voter turnout patterns: An information theory. *Public Choice*, 84(1-2):91–117, 1995.

[36] Thomas R Palfrey and Howard Rosenthal. A strategic calculus of voting. *Public Choice*, 41(1):7–53, 1983.

[37] Thomas R Palfrey and Howard Rosenthal. Voter participation and strategic uncertainty. *The American Political Science Review*, pages 62–78, 1985.

[38] Eric A Posner. *Law and social norms*. Harvard University Press, 2002.

[39] William H Riker and Peter C Ordeshook. A theory of the calculus of voting. *The American Political Science Review*, pages 25–42, 1968.

[40] William F Shughart II. Is voting rational? In *The Encyclopedia of Public Choice*, pages 650–653. Springer, 2004.

[41] Craig J Svoboda. Retrospective voting in gubernatorial elections: 1982 and 1986. *Political Research Quarterly*, 48(1):135–150, 1995.

[42] Edward R Tufte. Determinants of the outcomes of midterm congressional elections. *American Political Science Review*, 69(03):812–826, 1975.

[43] Jean-Robert Tyran. Voting when money and morals conflict: an experimental test of expressive voting. *Journal of Public Economics*, 88(7):1645–1664, 2004.