The Price of Deception in Spatial Social Choice

James P. Bailey and Craig A. Tovey

Abstract

Except for a few strategy-proof mechanisms on the real line, spatial social choice mechanisms are usually manipulable. But is it wise to treat all manipulations as equally bad? We introduce a new measure, the *price of deception*, to make finer distinctions than between "strategy-proof" and "manipulable." The price of deception, akin to but distinct from the price of anarchy in computer science, measures how much strategic behavior can alter the cost of the social choice. We propose it should be one of the criteria by which a selection rule is assessed. Supported by experimental economics data, our measure employs a novel *minimal dishonesty* criterion to refine the set of Nash equilibria. We calculate the price of deception for standard spatial selection rules, including 1-Median and 1-Mean, and find significant differences among them. We also find that a mechanism designer can significantly lessen the impact of manipulation by altering the set of allowed points to a hyper-rectangle. The concepts herein could be applied to other social choice scenarios in which a publicly known mechanism relies on private information.

1 Introduction

In spatial social choice each voter and each potential candidate is represented by a point in $\mathbb{R}^k$. The cost of a candidate to a voter is the Euclidean or some other metric distance between the two points. Spatial models are widely applied to facility location [21], preference aggregation [1, 10], political voting and policy selection [11, 8, 26, 33]. We denote the most preferred or *ideal point* of voter v as π_v.

A 1-Median selection mechanism selects a point that minimizes the total distance (L_1 norm) to the voters' ideal points. Both 1-Median selection and its generalization, the p-Median (which selects p points), have been carefully investigated for strategy-proof variants [29, 15, 9, 16]. But there has not been much work towards understanding the quality of the outcomes obtained when the selection rule is manipulable.

We introduce a measure of how much strategic behavior impacts the quality of the outcome of a selection rule. We call this measure the *price of deception.* To illustrate the idea, consider the 1-Median selection rule. Let the *sincere distance* of a point x be the total distance from the individuals' sincere ideal points to x. Let a *sincere optimal point* be a point with the smallest sincere distance. Suppose that the sincere distance of a point selected when individuals are strategic is at most 3 times the sincere distance of a sincere optimal point. Then the price of deception of the 1-Median problem would be at most 3. This kind of measure, the worst-case ratio of actual to best possible cost, has long been standard in computer science. In particular, the price of deception has a flavor akin to the price of anarchy [12, 3, 25] because both involve equilibria. However, they are not equivalent [2].

We propose that the price of deception should be one of the criteria by which a spatial social choice rule is assessed. In general, a rule can be cast as a minimizer of a social cost; i.e. in the 1-Median problem, the selected point minimizes the distance to the ideal points. If a point's cost has an intrinsic correspondence to its quality, then the price of deception has a correspondence to the capability of a selection rule to select a quality point. More generally, the price of deception measures a selection rule's ability to provide the expected sincere outcome despite fictitious play. Like the computational complexity of manipulation [23, 34], the price of deception offers finer distinctions than simply between "manipulable" and "non-manipulable."

The concept of the price of deception applies not only to spatial social choice, but to much of social choice in general. A centralized mechanism makes a decision that optimizes a measure of social benefit or cost based on information submitted by individuals. However, individuals have their own valuation of each possible outcome. Therefore they place a *game of deception* in which they provide possibly untruthful information, and experience outcomes in accordance with their own true valuation of the centralized decision made based on the information they provide.

We remark that the revelation principle [35, 36] is irrelevant to the price of deception. This is because revelation elicits sincere information only by yielding the same outcome that strategic information would yield. The revelation principle can be a powerful tool for analyzing outcomes. But for our purposes, the elicitation of sincere preference information is not an end in itself.

1.1 The Minimal Dishonesty Refinement

Analyzing "games of deception" is by no means novel [18, 42, 17, 41, 20, 40, 29, 5]. Yet the solution concept is neglected in many areas as researchers still simply label mechanisms as "manipulable" or "strategy-proof". While we make no attempt to describe the beliefs or intentions of other researchers, we believe there is a good reason to neglect the Nash equilibrium solution concept as we have described it, because the outcomes of games of deception often make little sense.

For instance, consider the 1-Median problem in $\mathbb{R}$ with an odd number $n \geq 3$ of voters. The selected point will coincide with the median voter's ideal point. Suppose every voter most prefers the point x. Then obviously, the selected point should be x. However, the Nash equilibrium solution concept tells us something different. If everyone lies and indicates that they prefer $y \neq x$, then the selected point will be y. Furthermore, if any single individual alters their submitted preferences then the median voter will still be located at y and the selected point will not change. Therefore for every possible point y, there is a Nash equilibrium where y is selected. *We believe these absurd equilibria are the biggest obstacle to understanding the effects of strategic behavior.*

To overcome this obstacle, we introduce a minimal dishonesty refinement to the Strategic Spatial Social Choice Game. An individual is minimally dishonest if any attempt to be more honest results in a strictly worse outcome.

We argue minimal dishonesty has an intuitive explanation, has a precedent in the voting community, show it is consistent with the current literature on strategy-proofness and cite a large amount of experimental evidence backing our refinement. If individual v can be more honest and get at least as good a result, then the individual would do so because lying causes guilt and because being somewhat truthful is cognitively easier than lying spuriously. Thus, there is some "utility" associated with being more honest. Therefore an individual that is not minimally dishonest is not acting in their best interest. Hence, we view minimally dishonest Nash equilibria as the set of reasonable outcomes in the Strategic Spatial Social Choice Game.

The voting literature sets a precedent for an honesty refinement with the partial honesty [13, 14, 24, 27, 37, 31] or truth-bias [39] refinement. This refinement assumes an individual is completely honest unless it negatively impacts their valuation of the outcome. However, this refinement assumes individuals evaluate honesty in a binary sense whereas later we learn individuals have a more nuanced idea of honesty. While this refinement does eliminate some of the absurd equilibria, it fails to eliminate some equilibria where individuals lie spuriously.

There is however a recent variant of truth-bias [38] where individual utilities are penalized by the size of the lie they tell. If individuals select their preference from a finite set, this variant can be shown to be equivalent to the minimal dishonesty refinement. However, in the setting of spatial social choice, these distorted costs can result in outcomes where individuals behave irrationally. We show both results in Appendix D.

The assumption of minimal dishonesty is also consistent with the current literature's assumptions of "strategy-proof" or "non-manipulable" mechanisms. Current literature assumes that if a mechanism is strategy-proof then every individual will be honest. However, strategy-proofness only requires that the sincere profile Π be at least one of the Nash equilibria. It does not require Π to be the only Nash equilibrium. The set of Nash equilibria predicts the outcome of events, yet the literature ignores all other equilibria when the mechanism is strategy-proof. This is reasonable because it makes little sense for people to lie when there is not an incentive to do so. Minimal dishonesty captures this behavior; in Section 2.2 we show that a mechanism is strategy-proof if and only if the sincere Π is the only minimally dishonest equilibrium.

We've argued that minimal dishonesty is logically intuitive, has a precedent in the voting literature, and that it explains the assumptions researchers make when using strategy-proof mechanisms. Most importantly, our hypothesis is supported by a substantial body of empirical evidence from the experimental economics and psychology literatures that people are averse to lying. Gneezy [19] experimentally finds that people do not lie unless there is a benefit. Hurkens and Kartik [22] perform additional experiments that confirm an aversion to lying, and show their and Gneezy's data to be consistent with the assumption that some people never lie and others always lie if it is to their benefit. Charness and Dufwenberg [6] experimentally find an aversion to lying and show that it is consistent with guilt avoidance. Battigalli *et al.* [4] experimentally find some contexts in which guilt is insufficient to explain aversion to deception. Several papers report evidence of a "pure," i.e., context-independent, aversion to lying [28, 7, 19, 22] that is significant but not sufficient to fully explain experimental data.

The set of research results we have cited here is by no means exhaustive. Two others are of particular relevance to our concept of minimal dishonesty. Mazar *et al.* [32] find that "people behave dishonestly enough to profit but honestly enough to delude themselves of their own integrity." Lundquist *et al.* [30] find that people have an aversion to lying which increases with the size of the lie. Both of these studies support our hypothesis that people will not lie more than is necessary to achieve a desirable outcome.

Some experimental evidence is less confirmatory of our hypothesis. Several studies, beginning with [19], have found an aversion to lying if doing so would disbenefit someone else substantially more than the benefit one would accrue.

1.2 Results

We begin by analyzing the Spatial Social Choice problem in $\mathbb{R}$ – that is where the set of points can be represented by a line segment. We analyze deterministic and random variants for the 1-Median problem and the 1-Mean problem and find interesting differences between them. We first show that the deterministic version of the 1-Median problem has a price of deception of 1. This is the best possible outcome we can hope for in a manipulable decision mechanism; while the 1-Median problem is manipulable, manipulation has no impact on the social cost. Next we show that the 1-Median problem with random tie-breaking has a price of deception of $\sqrt{2}$ – manipulation can cause social cost to increase by a factor of $\sqrt{2}$. This result is striking; a small change to a decision mechanism such as a tie-breaking rule may have significant effects on the impact of manipulation. Finally we analyze the 1-Mean problem with n players and show that the price of deception is between n and $2n$.

Next we examine the Spatial Social Choice problem in $\mathbb{R}^k$ for $k \geq 2$. Unlike the problem in $\mathbb{R}$, the set of possible points can take on a variety of shapes. We show that the shape of the set of points can have significant effects on the impact of manipulation. Specifically, if there are no restrictions on the set of points then the price of deception can be arbitrarily high. The mechanism designer can ameliorate the impact of manipulation by changing the set of allowed points to a hyperrectangle. In this setting we show the price of deception of the deterministic 1-Median, random 1-Median, and 1-Mean problem with n players to be 1, $\sqrt{2}$, and between n and $2n$ respectively. We also analyze the selection rule that minimizes the L_2 norm between the point and player's ideal points[1] and show that the price of deception is ∞ regardless of the set of allowed points.

2 Definitions

An instance of the Spatial Social Choice problem consists of a compact convex set $\mathcal{X} \subseteq \mathbb{R}^k$ of feasible points and a set V of individuals. Each individual $v \in V$ has an ideal point $\pi_v \in \mathcal{X}$ representing v's preferred point. Denote by Π the collection of π_v over all $v \in V$. $\Pi \in \mathcal{X}^{|V|}$ is called the *preference profile*. The profile Π is submitted to a selection rule r. The outcome $r(\Pi)$ corresponds to a point or a distribution of points.

The selection rule $r(\Pi)$ is an optimizer of a social cost function. We consider three mechanisms, each having a different social cost function:

$$C(\Pi, x) = \sum_{v \in V} ||\pi_v - x||_1 \qquad \text{(1-Median)}$$

$$C(\Pi, x) = \sum_{v \in V} ||\pi_v - x||_2^2 \qquad \text{(1-Mean)}$$

$$C(\Pi, x) = \sum_{v \in V} ||\pi_v - x||_2 \qquad (L_2 \text{ Norm})$$

Only the 1-Mean problem is guaranteed a unique optimizer. For the other problems we consider both deterministic and random tie-breaking rules.

Individual v's cost of the point $x \in \mathcal{X}$ is given by $c_v(\pi_v, x) = ||\pi_v - x||_{p_v} = (\sum_i (\pi_{vi} - x_i)^{p_v})^{1/p_v}$ for some $p_v \in (0, \infty)$. Typically p_v is taken to be 2 (the Euclidean norm). We derive our results for any p_v. In the event that the procedure r is random, we assume that individuals are risk-neutral and that $c_v(\pi_v, X)$ for some distribution X is equal to its expected value.

Strategic Spatial Social Choice Game

- Each individual v has an ideal point $\pi_v \in \mathcal{X}$. The collection of all ideal points is the (sincere) profile $\Pi = \{\pi_v\}_{v \in V}$.
- To play the game, individual v submits a putative ideal point $\bar{\pi}_v \in \mathcal{X}$. The collection of all submitted data is denoted $\bar{\Pi} = \{\bar{\pi}_v\}_{v \in V}$.
- It is common knowledge that a central decision mechanism will select point or distribution of points $r(\bar{\Pi})$ when given input $\bar{\Pi}$.
- Individual v evaluates $r(\bar{\Pi})$ according to v's sincere preferences π_v. Specifically, individual v's cost of the distribution of outcomes $r(\bar{\Pi})$ is $c_v(\pi_v, r(\bar{\Pi})) = E(||\pi_v - r(\bar{\Pi})||_{p_v})$ for some $p_v \in (0, \infty)$.

[1] In $\mathbb{R}$ minimizing the L_2 norm is equivalent to solving the 1-Median problem.

A set of putative preferences $\bar{\Pi}$ forms a pure strategy Nash equilibrium if no individual v would obtain an outcome they sincerely prefer to $r(\bar{\Pi})$ (with respect to π_v) by altering $\bar{\pi}_v$. Formally, $\bar{\Pi}$ is a Nash equilibrium if and only if

$$c_v(\bar{\Pi}) \leq c_v([\bar{\Pi}_{-v}, \bar{\pi}'_v]) \text{ for } \bar{\pi}'_v \in \mathcal{X} \text{ and } v \in V \qquad \text{(Nash equilibrium)}$$

where $[\bar{\Pi}_{-v}, \bar{\pi}'_v]$ denotes the profile obtained from $\bar{\Pi}$ by replacing $\bar{\pi}_v$ with $\bar{\pi}'_v$.

Example 2.1. A Nash Equilibrium of the Strategic Spatial Social Choice Game.

Consider the 1-Mean problem. It is well known that the optimizer of (1-Mean) is $r(\Pi) = \sum_{v \in V} \frac{\pi_i}{|V|}$. *Consider the feasible region* $\mathcal{X} = \{x \in \mathbb{R}^2 : (0,0) \leq x \leq (1,1)\}$ *and the sincere preferences* $\pi_1 = (0,0), \pi_2 = (0, \frac{1}{3})$ *and* $\pi_3 = (\frac{1}{3}, 0)$. *With respect to these preferences, the selected point is located at* $r(\Pi) = (\frac{1}{9}, \frac{1}{9})$. *This corresponds to a cost of* $C(\Pi, r(\Pi)) = \sum_{i=1}^{3} ||\pi_i - r(\Pi)||_2^2 = \frac{4}{27}$. *The region* $\mathcal{X}$ *and sincere preferences,* Π, *are given in Figure 1.*

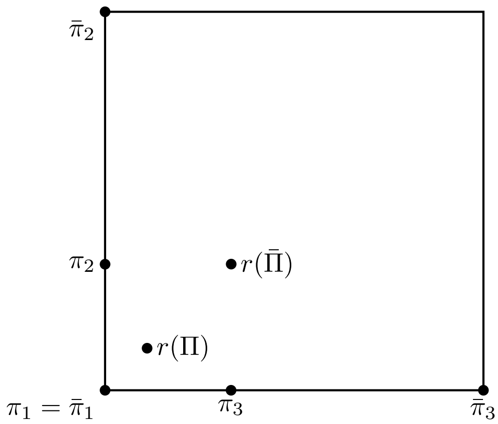

Figure 1: Sincere and Putative Preferences for Example 2.1

Individual 1 would like to move the point to the lower left, individual 2 to the upper left, and individual 3 to the lower right. A Nash equilibrium where individuals attempt to do just this is given by $\bar{\pi}_1 = (0,0), \bar{\pi}_2 = (0,1)$, *and* $\bar{\pi}_3 = (1,0)$. *With respect to the submitted preferences* $\bar{\Pi}$, *the point is* $r(\bar{\Pi}) = (\frac{1}{3}, \frac{1}{3})$. *The central decision mechanism believes it has selected a point with cost* $C(\bar{\Pi}, r(\bar{\Pi})) = \sum_{i=1}^{3} ||\bar{\pi}_i - r(\bar{\Pi})||_2^2 = \frac{4}{3}$. *However, with respect to the true preferences* Π, *the point actually costs* $C(\Pi, r(\bar{\Pi})) = \sum_{i=1}^{3} ||\pi_i - r(\bar{\Pi})||_2^2 = \frac{4}{9}$.

To see that $\bar{\Pi}$ *corresponds to a Nash equilibrium, first consider individual 1. If individual 1 alters her submitted ideal point* $\bar{\pi}_1$, *then she must move it up or to the right. Such an action causes the point to move up or to the right respectively. Both possibilities cause the point to move further away from* π_1 *and therefore individual* 1 *cannot alter her submitted preferences to get a better result. Individual 2 and Individual 3 are giving best responses by symmetric reasoning and* $\bar{\Pi}$ *is a pure strategy Nash equilibrium.*

Example 2.1 demonstrates that manipulation in the 1-Mean problem can cause the social cost to increase from $\frac{4}{27}$ to $\frac{4}{9}$ – a cost that is 3 times worse.

2.1 The Price of Deception

The price of deception is a worst-case analysis of how much manipulation can impact the social cost.

Definition 2.2. Let r be a selection rule that minimizes the real-valued cost function C. Let $NE(\Pi)$ denote the set of equilibria of the Strategic Spatial Social Choice Game with procedure r given the sincere profile Π. Then the price of deception of r is

$$\sup_{\Pi \in \mathcal{P}} \sup_{\bar{\Pi} \in NE(\Pi)} \frac{E(C(\Pi, r(\bar{\Pi})))}{C(\Pi, r(\Pi))} \qquad \text{(The Price of Deception)}$$

Proving that the price of deception is u requires two parts. First, there must be an instance (or family of instances) showing that manipulation can cause social cost to increase by a factor of u (arbitrarily close to u) indicating the price of deception is at least u. Second, there cannot be an instance where manipulation can cause social cost to increase by a factor more than u indicating the price of deception is at most u.

For instance, Example 2.1 demonstrated that in the 1-Mean problem with 3 individuals may result in the social cost increasing by a factor of 3 and therefore the price of deception is at least 3. To show a price of deception of 3 we would also have to provide a proof that manipulation cannot cause social cost to increase by a factor more than 3.

2.2 The Minimally Dishonest Refinement

The necessity of a refinement on the set of Nash equilibria was established in Section 1.1. We've also shown that our minimally dishonest refinement is intuitive, has a precedent in the voting literature, and is backed by a large amount of experimental evidence. In this section, we formally define the minimally dishonest refinement and show that it is consistent with the literature's assumption of honesty in strategy-proof mechanisms.

Definition 2.3. Let Π be the sincere preferences and let $\bar{\Pi}$ be a Nash equilibrium in the Strategic Spatial Social Choice Game. An individual v is minimally dishonest if $||\pi_v - \bar{\pi}'_v|| < ||\pi_v - \bar{\pi}_v||$ implies $c_v(\pi_v, r([\bar{\Pi}_{-v}, \bar{\pi}'_v)]) > c_v(\pi_v, r(\bar{\Pi}))$.

An individual is minimally dishonest if being more honest always results in a strictly worse outcome for the individual. A minimally dishonest Nash equilibrium is a Nash equilibrium where every individual is minimally dishonest.

Definition 2.4. A mechanism is strategy-proof if the sincere π_v is always a best response to Π_{-v} for all $v \in V$ and Π.

Theorem 2.5. *A mechanism is strategy-proof if and only if* Π *is the only minimally dishonest equilibrium for any sincere* Π.

Proof. For the first direction, since the mechanism is strategy-proof, it is always a best response for individual v to submit the honest π_v regardless of all other preferences. Therefore π_v is the unique minimally dishonest best response for individual v. This holds for all v and therefore Π is the unique minimally dishonest Nash equilibrium.

For the second direction, let Π be an arbitrary set of sincere preferences. Since Π is a minimally dishonest equilibrium for the sincere profile Π, π_v is a best response to Π_{-v} for all v. This holds for all π_v and Π_{-v} and therefore honesty is always a best response and the mechanism is strategy-proof. □

3 Prices of Deception for Selection Rules in $\mathbb{R}$

3.1 The 1-Median Problem with Deterministic Tie-Breaking

We begin by breaking ties deterministically. Let $a(\Pi)_i = \text{argmin}\{\pi_{vi} : |\{v' : \pi_{v'i} \leq \pi_{vi}\}| \geq |V|/2\}$ and $b(\Pi)_i = \text{argmax}\{\pi_{vi} : |\{v' : \pi_{v'i} \geq \pi_{vi}\}| \geq |V|/2\}$. It is now straightforward to verify that x minimizes (1-Median) if and only if $x \in [a(\Pi), b(\Pi)]$. Using this notation, we can now define the λ-1-Median problem.

Definition 3.1. Let $\lambda \in [0,1]^k$. For the profile Π, let $\{x \in \mathbb{R}^k : a(\Pi) \leq x \leq b(\Pi)\}$ be the set of optimal points in the 1-Median problem. The λ-1-Median problem selects point $r(\Pi) = (\mathbf{1} - \lambda)a(\Pi) + \lambda b(\Pi)$.

Theorem 3.2. *When individuals are minimally dishonest, the price of deception for the λ-1-Median problem in $\mathbb{R}$ is* 1*.*

Proof of Theorem 3.2. Let Π and $\bar{\Pi}$ be sincere and submitted preferences where $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. It suffices to show $r(\bar{\Pi}) = (1-\lambda)a(\bar{\Pi}) + \lambda b(\bar{\Pi}) \in [a(\Pi), b(\Pi)]$ since every $x \in [a(\Pi), b(\Pi)]$ is an optimal solution to the 1-Median problem for preferences Π. For contradiction and without loss of generality, suppose $(1-\lambda)a(\bar{\Pi}) + \lambda b(\bar{\Pi}) < a(\Pi)$.

First we claim that if $\pi_v \geq b(\Pi)$ then $\bar{\pi}_v \geq b(\Pi)$. For contradiction suppose there exists a v where $\pi_v \geq b(\Pi)$ but $\bar{\pi}_v < b(\Pi)$. If v instead submits $\bar{\pi}'_v = \bar{\pi}_v + \epsilon$ for some $\epsilon > 0$ then $r([\bar{\Pi}_{-v}, \bar{\pi}'_v]) \geq r(\bar{\Pi})$. Moreover, (1-Median) is continuous and therefore ϵ can be selected sufficiently small so that $r(\bar{\Pi}) \leq r([\bar{\Pi}_{-v}, \bar{\pi}'_v]) \leq \pi_v$. Therefore, v can submit the more honest $\bar{\pi}'_v$ and obtain at least as good an outcome contradicting minimal dishonesty. Therefore the claim holds.

By our claim, $|\{v : \bar{\pi}_v \geq b(\Pi)\}| \geq |\{v : \pi_v \geq b(\Pi)\}| \geq \frac{|V|}{2}$. If $|\{v : \pi_v \geq b(\Pi)\}| > \frac{|V|}{2}$ then $|\{v : \bar{\pi}_v < b(\Pi)\}| < \frac{|V|}{2}$ implying $a(\bar{\Pi}) \geq b(\Pi)$, a contradiction. Therefore $|\{v : \pi_v \geq b(\Pi)\}| = \frac{|V|}{2}$ implying $|V|$ is even. This completes the proof when $|V|$ is odd and implies $a(\Pi) < b(\Pi)$ when $|V|$ is even.

Let v' be such that $\pi_{v'} = a(\Pi)$. If $\bar{\pi}_v \geq a(\Pi)$ then $|\{v : \bar{\pi}_v \geq a(\Pi)\}| \geq |\{v : \pi_v \geq b(\Pi)\}| + 1 \geq \frac{n}{2} + 1$ implying $|\{v : \bar{\pi}_v < a(\Pi)\}| < \frac{n}{2}$ and $a(\bar{\Pi}) \geq a(\Pi)$, a contradiction. Therefore $\bar{\pi}_v < a(\Pi)$. Now suppose instead v' submits $\bar{\pi}'_v = \bar{\pi}_{v'} + \epsilon$ for some $\epsilon > 0$. Similar to before, ϵ can be selected sufficiently small so that $r(\bar{\Pi}) \leq r([\bar{\Pi}_{-v'}, \bar{\pi}'_{v'}]) \leq \pi_{v'}$. Therefore, v' can submit the more honest $\bar{\pi}'_{v'}$ and obtain a least as good of an outcome contradicting minimal dishonesty completing the proof of the theorem. □

Theorem 3.2 shows that manipulation may alter the outcome of the decision mechanism but manipulation will not impact the social cost. Furthermore, we can select λ such that the mechanism is not manipulable. Specifically, for $\lambda \in \{0,1\}$, the λ-1-Median problem is strategy-proof.

Theorem 3.3. *Fix $\lambda \in \{0,1\}$. The λ-1-Median problem in $\mathbb{R}$ is strategy-proof.*

Proof. By definition of strategy-proof, we need to show that the sincere π_v is a best response to Π_{-v} for all v and Π_{-v}. By symmetry, assume $\lambda = 1$ and $r(\Pi) = b(\Pi)$. If $\pi_v = b(\Pi)$ then it trivially a best response for v to be honest. If $\pi_v > b(\Pi)$, then v can only change the outcome by submitting $\bar{\pi}_v < b(\Pi)$ resulting in $r([\Pi_{-v}, \bar{\pi}_v]) \leq r(\Pi) < \pi_v$. However, this solution is no better for v and therefore π_v is a best response. Symmetrically, π_v is a best response if $\pi_v < b(\Pi)$ completing the proof of the theorem. □

When $|V|$ is odd, $a(\bar{\Pi}) = b(\bar{\Pi})$ and the tie-breaking rule is never invoked. Therefore the 1-Median problem is strategy-proof with an odd number of voters.

Corollary 3.4. *When $|V|$ is odd the 1-Median problem in $\mathbb{R}$ is strategy-proof.*

3.2 The 1-Median Problem with Random Tie-Breaking

Next, we consider breaking ties uniformly at random. If $|V|$ is odd then there are no ties and the 1-Median problem is strategy-proof. When there is an even number of voters we show that the price of deception is $\sqrt{2}$. First we characterize an individual's best response.

Lemma 3.5. *Suppose $|V| = 2$ and each voter must submit a location in the interval $\mathcal{X} = [l, u]$ for the 1-Median problem when breaking ties uniformly at random. If $\bar{\pi}_1 \leq \pi_2$, then voter 2's unique minimally dishonest best response is $\bar{\pi}_2 = min\{u, \bar{\pi}_1 + \sqrt{2}(\pi_2 - \bar{\pi}_1)\}$.*

The proof of Lemma 3.5 is deferred to Appendix A.

Theorem 3.6. *Let $|V| = 2$. When both individuals are minimally dishonest and ties are broken uniformly at random, the price of deception for the 1-Median problem in $\mathbb{R}$ is $\sqrt{2}$.*

Proof. Without loss of generality $\pi_1 \leq \pi_2$. First we consider $\pi_1 = \pi_2$. Following from Lemma 3.5, the only minimally dishonest Nash equilibrium is $\bar{\Pi} = \Pi$. In this setting manipulation has no impact on social cost and we assume $\pi_1 < \pi_2$.

By scaling, translating, and reflecting we may assume $\pi_1 = -1$ and $\pi_2 = 1$ and that $u \geq -l$. The sincere outcome $r(\bar{\Pi})$ selects a point uniformly at random in $[-1, 1]$ with cost $U(\Pi, r(\Pi)) = 2$. Again let $\bar{\Pi}$ be a minimally dishonest Nash equilibrium. By Lemma A.1 of Appendix A, $\bar{\pi}_1 \leq \pi_2$. Therefore by Lemma 3.5 $\bar{\pi}_2 = min\{u, \bar{\pi}_1 + \sqrt{2}(1 - \bar{\pi}_1)\}$. Symmetrically, $\bar{\pi}_1 = max\{l, \bar{\pi}_2 + \sqrt{2}(-1 - \bar{\pi}_2)\}$.

Observe that if $\bar{\pi}_2 = u$ then

$$\bar{\pi}_2 + \sqrt{2}(-1 - \bar{\pi}_2) = (1 - \sqrt{2})u - \sqrt{2} \tag{1}$$

$$\leq (\sqrt{2} - 1)l - \sqrt{2} < l \tag{2}$$

This implies that if $\bar{\pi}_2 = u$ then $\bar{\pi}_1 = l$. Therefore we can break the problem into two cases: $\bar{\pi}_1 > l$, $\bar{\pi}_2 < u$; and $\bar{\pi}_1 = l$.

Case 1: $\bar{\pi}_1 > l$, $\bar{\pi}_2 < u$. For brevity, let $C = C(\Pi, r(\bar{\Pi}))$. Therefore

$$E(C) = \frac{\bar{\pi}_2^2 + \bar{\pi}_1^2 + 2}{\bar{\pi}_2 - \bar{\pi}_1}. \tag{3}$$

By Lemma 3.5

$$\bar{\pi}_2 = (1 - \sqrt{2})\bar{\pi}_1 + \sqrt{2} \tag{4}$$

$$\bar{\pi}_1 = (1 - \sqrt{2})\bar{\pi}_2 - \sqrt{2} \tag{5}$$

which has the unique solution $\bar{\pi}_2 = -\bar{\pi}_1 = 1 + \sqrt{2}$ yielding $E(C) = 2\sqrt{2}$. When everyone is sincere $C(\Pi, r(\Pi)) = 2$ and the price of deception is at most $\sqrt{2}$.

Case 2: $\bar{\pi}_1 = l$. By Lemma 3.5, $l \geq (1 - \sqrt{2})\bar{\pi}_2 - \sqrt{2}$ and $\bar{\pi}_2 \leq (1 - \sqrt{2})l + \sqrt{2}$. Combining both inequalities, $l \geq -\sqrt{2} - 1$. Combining the last two inequalities, $\bar{\pi}_2 \leq \sqrt{2} + 1$.

The function $(z^2 + y^2 + 2)/(z - y)$ is increasing on the interval $0 \leq z \leq \sqrt{2} + 1$ when $y \in [-\sqrt{2} - 1, 0]$ and decreasing on the interval $-\sqrt{2} - 1 \leq y \leq 0$ when $z \in [0, \sqrt{2} + 1]$. Therefore

$$E(C) = \frac{\bar{\pi}_2^2 + \bar{\pi}_1^2 + 2}{\bar{\pi}_2 - \bar{\pi}_1} \leq \frac{(\sqrt{2} + 1)^2 + (-\sqrt{2} - 1)^2 + 2}{(\sqrt{2} + 1) - (-\sqrt{2} - 1)} = 2\sqrt{2}. \tag{6}$$

As in Case 1, the cost of the sincere outcome is 2 and the price of deception is at most $\sqrt{2}$. An instance showing the price of deception is at least $\sqrt{2}$ follows from Case 1 with u and l sufficiently large completing the proof of the theorem. □

Theorem 3.7. *Suppose $|V|$ is even. When individuals are minimally dishonest and ties are broken uniformly at random, the price of deception for the 1-Median problem in $\mathbb{R}$ is $\sqrt{2}$.*

We prove Theorem 3.7 by ordering the individuals so that $\pi_{-n} \leq \pi_{-n+1} \leq ... \leq \pi_{-1} \leq \pi_1 \leq \pi_2 \leq ...\pi_n$. We then reduce the $|V|$-player game to a 2-player game between players -1 and 1 and show manipulation impacts social cost in the 2-player game at least as much as it impacts social cost in the $|V|$-player game. Theorem 3.7 then follows from Theorem 3.6. The full details of the proof can be found in Appendix A.

3.3 The 1-Mean Problem

Theorem 3.8. *The price of deception of the 1-Mean problem in $\mathbb{R}$ is between $|V|$ and $2|V|$.*

Proof. First we show an upper bound of $2|V|$. Let $\bar{\Pi}$ be a Nash equilibrium for the sincere profile Π. Let $[a,b]$ be the smallest interval such that $\Pi \in [a,b]$. Without loss of generality we may assume $a = 0$. As established, the decision mechanism selects $r(\Pi) = \sum_{v \in V} \frac{\pi_v}{|V|}$ to minimize $\sum ||\pi_i - x||_2^2$ and therefore $r_j(\Pi) \in [0,b]$. We then have

$$\sum_{v \in V} ||\pi_v - r(\Pi)||_2^2 = \sum_{v \in V} (\pi_v - r(\Pi))^2 \tag{7}$$

$$\geq \min_{v \in V} (\pi_v - r(\Pi))^2 + \max_{v \in V} (\pi_v - r(\Pi))^2 \tag{8}$$

$$= r(\Pi)^2 + (b - r(\Pi))^2 \tag{9}$$

$$\geq \left(\frac{b}{2}\right)^2 + \left(b - \frac{b}{2}\right)^2 = \frac{b^2}{2} \tag{10}$$

and therefore the sincere cost of the point is at least $\frac{b^2}{2}$.

Next we claim $r(\bar{\Pi}) \in [0,b]$. Suppose instead $r(\bar{\Pi}) < 0$ and there exists an individual v that submits $\bar{\pi}_v \leq r(\bar{\Pi}) < 0$. If v instead submits the location 0, then the selected point moves closer to 0, a strict improvement for every individual contradicting that $\bar{\Pi}$ is a Nash equilibrium. Therefore $r(\bar{\Pi}) \geq 0$. Symmetrically $r(\bar{\Pi}) \leq b$ completing the claim.

Since $r(\bar{\Pi}) \in [0,b]$ and $\pi_v \in [0,b]$ for all v, the sincere cost for the point is

$$\sum_{v \in V} ||\pi_v - r(\bar{\Pi})||_2^2 = \sum_{v \in V} (\pi_v - r(\bar{\Pi}))^2 \tag{11}$$

$$\leq \sum_{v \in V} b^2 = |V| b^2. \tag{12}$$

Therefore, the manipulation causes the social cost to increase by a factor of

$$\frac{\sum_{v \in V} ||\bar{\pi}_i - r(\Pi)||_2^2}{\sum_{v \in V} ||\pi_i - r(\Pi)||_2^2} \leq \frac{|V| b^2}{\frac{b^2}{2}} = 2|V| \tag{13}$$

We now present an instance with a price of deception of $|V|$. Let $\mathcal{X} = [0, |V|]$. Let $\pi_1 = 1$ and $\pi_v = 0$ for all $v \in V \setminus \{1\}$. If everyone is sincere, then the selected point is $r(\Pi) = \frac{1}{|V|}$ with a social cost of $\frac{|V|-1}{|V|}$.

Now consider the putative preferences $\bar{\Pi}$ where $\bar{\pi}_1 = |V|$ and $\bar{\pi}_v = \pi_v$ for all other v. With respect to these preferences, the selected point is $r(\bar{\Pi}) = 1$ for a sincere social cost of $|V| - 1$. It is straightforward to verify that $\bar{\Pi}$ is a minimally dishonest Nash equilibrium and therefore manipulation can cause the social cost to increase by a factor of $|V|$. Thus, the price of deception is between $|V|$ and $2|V|$ completing the proof of the theorem. □

4 Prices of Deception for Selection Rules in $\mathbb{R}^k$ for $k \geq 2$

4.1 The 1-Median Problem with Deterministic Tie-Breaking

Since the 1-Median problem is separable we might expect that Theorems 3.2 and 3.7 hold in higher dimensions. However, while determining $r(\Pi)$ might be separable, determining a minimally dishonest best response $\bar{\pi}_v$ is not. Regrettably, the lack of separability in determining a best response can lead to a large price of deception in higher dimensions.

Theorem 4.1. *Suppose $\lambda_1 \in (0,1)^k$ and $|V|$ is even. When individuals are minimally dishonest, the price of deception of the λ-1-Median problem in $\mathbb{R}^k$ is ∞ for $k \geq 2$.*

Proof. It suffices to show the result for $k = 2$ since the bad example can always be placed in higher dimensions. Let $|V| = 2n$ for some integer $n \geq 2$. The sincere preferences are $\pi_v = 0$ for all v. The set of feasible points is $\mathcal{X} = conv.hull(\pi_1, a, b)$ as shown in Figure 2 where $conv.hull(S)$ is the convex hull of S.

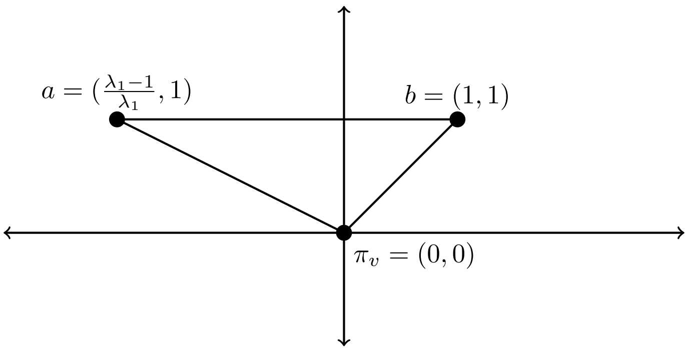

Figure 2: Preferences Showing the Price of Deception is ∞ for the λ-1-Median Problem.

If everyone is honest, then the selected point is $r(\Pi) = (0,0)$ with total cost 0. Suppose half the voters submit $\bar{\pi}_v = a$ and half the voters submit $\bar{\pi}_v = b$. Then $r(\bar{\Pi}) = \lambda a + (\mathbf{1} - \lambda) b = (0,1)$ with a sincere cost of $|V|$. If $\bar{\Pi}$ is a minimally dishonest Nash equilibrium, then the price of deception is ∞.

We now show that $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. Consider voter v that submits $\bar{\pi}_v = b$. Regardless of how she alters her submitted information, more than half the voters submit a height of 1 and therefore she cannot change the height of the point. Moreover, if she alters her preferences at all then the point moves to the left corresponding to a worse outcome for her. Therefore v is providing a minimally dishonest best response. Symmetrically voter v submitting $\bar{\pi}_v = a$ is providing a minimally dishonest best response and $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. □

Theorem 4.1 shows there are sets of possible locations such that the price of deception is arbitrarily large. As mechanism designers, we may alter $\mathcal{X}$ such that we obtain more desirable results. Specifically, we show that if $\mathcal{X} = \{x \in \mathbb{R}^k : l \leq x \leq u\}$ then manipulation does not impact social cost in the λ-1-Median problem.

Theorem 4.2. *Suppose $\mathcal{X} = \{x \in \mathbb{R}^k : l \leq x \leq u\}$. When individuals are minimally dishonest, the price of deception of the λ-1-Median problem is 1 for all $\lambda \in [0,1]^k$.*

Unlike the instance shown in Figure 2, selecting jth coordinate of $\bar{\pi}_v$ places no restrictions on the remainder of $\bar{\pi}_v$ and therefore the a minimally dishonest best response is a separable problem. Thus, the k-dimensional case reduces to the 1-dimensional case and Theorems 4.2

holds by Theorem 3.3. Moreover, we also immediately obtain a variety of strategy-proof mechanisms regardless of $\mathcal{X}$.

Theorem 4.3. *Fix $\lambda \in \{0,1\}^k$ and $\mathcal{X} = \{x \in \mathbb{R}^k : l \leq x \leq u\}$. The λ-1-Median problem is strategy-proof.*

Given the description of $\mathcal{X}$, determining a best response is again separable. Therefore it suffices to show the result holds for the one dimensional case and Theorem 4.3 follows from Theorem 3.3.

Theorem 4.4. *Fix $\lambda \in \{0,1\}^k$ and let $\mathcal{X}$ be arbitrary. The λ-1-Median problem is strategy-proof.*

Proof. We begin by defining a second voting procedure by expanding $\mathcal{X}$. Let $l, u \in \mathbb{R}^k$ be such that $\mathcal{X} \subseteq \mathcal{X}' = \{x \in \mathbb{R}^k : l \leq x \leq u\}$. Consider the λ-1-Median problem on the set $\mathcal{X}'$ with sincere preferences Π. By Theorem 4.3, $\pi_v \in \mathcal{X}$ is a best response for each voter v.

Since $\mathcal{X} \subseteq \mathcal{X}'$ and $\pi_v \in \mathcal{X}$ for all v, π_v must also be a best response in the original procedure with the smaller feasible set $\mathcal{X}$ completing the proof of the theorem. □

Corollary 4.5. *Let $\mathcal{X}$ be arbitrary. When $|V|$ is odd, the 1-Median problem is strategy-proof.*

As with Corollary 3.4, Corollary 4.5 immediately holds because there are never any ties when $|V|$ is odd.

4.2 The 1-Median Problem with Random Tie-Breaking

Theorem 4.6. *When individuals are minimally dishonest and ties are broken uniformly at random, the price of deception of the 1-Median problem in $\mathbb{R}^k$ is ∞ for $k \geq 2$.*

The proof of Theorem 4.6 is similar to the proof of Theorem 4.1 and is deferred to Appendix B. Thus once again the price of deception may be arbitrarily high when no restrictions are placed on the set of allowed points $\mathcal{X}$.

Theorem 4.7. *Suppose $|V|$ is even and $\mathcal{X} = \{x \in \mathbb{R}^k : l \leq x \leq u\}$. When individuals are minimally dishonest and ties are broken uniformly at random, the price of deception of the 1-Median problem is $\sqrt{2}$.*

Finding a minimally dishonest best response is again a separable problem and Theorem 4.7 follows directly from Theorem 3.7.

4.3 The 1-Mean Problem

Theorem 4.8. *The price of deception of the 1-Mean problem in $\mathbb{R}^k$ for $k \geq 2$ is ∞.*

The proof of Theorem 4.8 is similar to the proofs of Theorems 4.1 and 4.6 and is deferred to Appendix C.

Theorem 4.9. *Suppose the set of feasible points is $\mathcal{X} = \{x \in \mathbb{R}^k : l \leq x \leq u\}$. Then the price of deception of of the 1-Mean problem is between $|V|$ and $2|V|$.*

The proof of Theorem 4.9 follows identically to Theorem 3.8.

4.4 Minimizing L_2 Norm

Theorem 4.10. *Let $\mathcal{X}$ be an arbitrary set that contains an open two-dimensional subspace. When individuals are minimally dishonest and ties are broken by selecting the center point or uniformly at random, the price of deception when minimizing the sum of L_2 norm distances is ∞.*

Proof. It suffices to show this is true in $\mathbb{R}^2$. By scaling and shifting $\mathcal{X}$ we may assume $\{(0,0),(-1,1),(1,1)\} \in \mathcal{X}$. Let $|V| \geq 4$ be even and suppose $\pi_v = (0,0)$ for all v. If everyone is honest then the selected point is $r(\Pi) = (0,0)$ with a social cost of 0. Suppose instead that the players submit $\bar{\Pi}$ where half the voters submit $(-1,1)$ and the other half submits $(1,1)$. With respect to these preferences, the point is either $(0,1)$ or selected uniformly at random between $(-1,1)$ and $(1,1)$ with a sincere cost of at least 1. Furthermore, if $\bar{\Pi}$ is a minimally dishonest Nash equilibrium, then the price of deception of this instance is ∞.

We now show $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. It suffices to examine player v that submits $\bar{\pi}_v = (-1,1)$. If player v instead submits $\bar{\pi}'_v = (\bar{\pi}'_{v1}, \bar{\pi}'_{v2})$ where $\bar{\pi}'_{v2} \neq \bar{\pi}_{v2}$, then $r([\bar{\Pi}_{-1}, \bar{\pi}'_1]) = (1,1)$ yielding a worse outcome for player v. If $\bar{\pi}'_v$ is directly to the left of $\bar{\pi}_v$, then the outcome does not change and player v is less honest. If player v submits $\bar{\pi}'_v = (w,1)$ for some $w \in (-1,1)$, then the point is either $(\frac{1+w}{2} \geq 0, 1)$ or selected uniformly at random between $(w,1)$ and $(1,1)$. Both correspond to worse outcomes for player v. Finally, if player v submits $\bar{\pi}'_v = (w,1)$ for some $w \geq 1$, then $r([\bar{\Pi}_{-1}, \bar{\pi}'_1]) = (1,1)$ yielding a worse outcome for player v. All possibilities yield either worse outcomes for player v or cause player v to be less honest without any benefit. Thus player v is giving the unique minimally dishonest best response and $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. Therefore the price of deception when minimizing the L_2 norm is ∞. □

5 Conclusion

Our results show that strategic behavior impacts the outcomes of different decision mechanisms in very different ways. In the case of deterministic variants of the 1-Median problem, we show that while the selection rule is manipulable, strategic behavior has no impact on the quality of the outcome. However, if we make a small change to the 1-Median problem and break ties randomly instead, then the manipulation has a much larger impact on social cost. For the 1-Mean problem we see that the impact of manipulation scales linearly with the number of voters and when minimizing the L_2 norm manipulation can always lead to arbitrarily poor outcomes. We also show that the mechanism designer can ameliorate the impact of manipulation simply by altering the set of allowed outcomes. Importantly, the price of deception successfully captures this behavior and thus we recommend that it be one of the criteria in which a decision mechanism is assessed.

Acknowledgments. Work on the price of deception is supported by the United States government through the National Science Foundation Grant 1335301. James P. Bailey is funded by the Singapore government through the Ministry of Education AcRF Tier 2 Grant 2016-T2-1-170.

References

[1] Nir Ailon, Moses Charikar, and Alantha Newman. Aggregating inconsistent information: Ranking and clustering. *jacm*, 55(5):23:1–23:27, 2008.

[2] James P. Bailey. *The Price of Deception in Social Choice*. PhD thesis, Georgia Institute of Technology, 2017.

[3] John Bartholdi, Thomas Seeley, Craig Tovey, and John VandeVate. The pattern and effectiveness of forager allocation among food sources in honey bee colonies. *jtb*, 160:23–40, 1993.

[4] Pierpaolo Battigalli, Gary Charness, and Martin Dufwenberg. Deception: The role of guilt. *Journal of Economic Behavior & Organization*, 93:227–232, 2013.

[5] Simina Branzei, Ioannis Caragiannis, Jamie Morgenstern, and Ariel D. Procaccia. How bad is selfish voting? In Marie desJardins and Michael L. Littman, editors, *AAAI*. AAAI Press, 2013.

[6] Gary Charness and Martin Dufwenberg. Promises and partnership. *Econometrica*, 74(6):1579–1601, 2006.

[7] Gary Charness and Martin Dufwenberg. Bare promises: An experiment. *Economics Letters*, 107(2):281–283, 2010.

[8] O. Davis, M. Degroot, and M. Hinich. Social preference orderings and majority rule. *Econometrica*, 40:147–157, 1972.

[9] Elad Dokow and Dvir Falik. Models of manipulation on aggregation of binary evaluations. *CoRR*, abs/1201.6388, 2012. Presented at COMSOC 2012.

[10] Elad Dokow and Ron Holzman. Aggregation of binary evaluations. *Journal of Economic Theory*, 145(2):495 – 511, 2010. Judgment Aggregation.

[11] A. Downs. *An Economic Theory of Democracy*. Harper and Row, 1957.

[12] Pradeep Dubey. Inefficiency of nash equilibria. *Mathematics of Operations Research*, 11(1):1–8, 1986.

[13] Bhaskar Dutta and Jean-François Laslier. Costless honesty in voting. In *Proceedings of Social Choice and Welfare*, 2010.

[14] Bhaskar Dutta and Arunava Sen. Nash implementation with partially honest individuals. *Games and Economic Behavior*, 74(1):154 – 169, 2012.

[15] Bruno Escoffier, Laurent Gourvès, Nguyen Kim Thang, Fanny Pascual, and Olivier Spanjaard. Strategy-proof mechanisms for facility location games with many facilities. In *Proceedings of the Second International Conference on Algorithmic Decision Theory*, ADT'11, pages 67–81, Berlin, Heidelberg, 2011. Springer-Verlag.

[16] Dimitris Fotakis and Christos Tzamos. Strategy-proof facility location for concave cost functions. *CoRR*, abs/1305.3333, 2013.

[17] Peter Gärdenfors. Manipulation of social choice functions. *Journal of Economic Theory*, 13(2):217–228, 1976.

[18] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica: Journal of the Econometric Society*, pages 587–601, 1973.

[19] Uri Gneezy. Deception: The role of consequences. *The American Economic Review*, 95(1):384–394, 2005.

[20] Dan Gusfield and Robert Irving. *The Stable Marriage Problem: Structure and Algorithms*. The MIT Press, 1989.

[21] Harold Hotelling. Stability in competition. In *The Collected Economics Articles of Harold Hotelling*, pages 50–63. Springer, 1990.

[22] Sjaak Hurkens and Navin Kartik. Would i lie to you? on social preferences and lying aversion. *Experimental Economics*, 12(2):180–192, 2009.

[23] John J. Bartholdi III, Craig A. Tovey, and Michael A. Trick. The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6(3):227–241, 1989.

[24] Navin Kartik, Olivier Tercieux, and Richard Holden. Simple mechanisms and preferences for honesty. *Games and Economic Behavior*, 83(C):284–290, 2014.

[25] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413. Springer, 1999.

[26] G. Kramer. Sophisticated voting over multidimensional spaces. *Journal of Mathematical Sociology*, 2:165–180, 1972.

[27] Jean-François Laslier, Matías Núñez, and Carlos Pimienta. Reaching consensus through approval bargaining. *Games and Economic Behavior*, 104:241 – 251, 2017.

[28] Raúl López-Pérez and Eli Spiegelman. Why do people tell the truth? experimental evidence for pure lie aversion. *Experimental Economics*, 16(3):233–247, 2013.

[29] Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. Asymptotically optimal strategy-proof mechanisms for two-facility games. In *Proceedings of the 11th ACM Conference on Electronic Commerce*, EC '10, pages 315–324, New York, NY, USA, 2010. ACM.

[30] Tobias Lundquist, Tore Ellingsen, Erik Gribbe, and Magnus Johannesson. The aversion to lying. *Journal of Economic Behavior & Organization*, 70(1):81–92, 2009.

[31] Hitoshi Matsushima. Role of honesty in full implementation. *Journal of Economic Theory*, 139(1):353 – 359, 2008.

[32] Nina Mazar, On Amir, and Dan Ariely. The dishonesty of honest people: A theory of self-concept maintenance. *Journal of Marketing Research*, pages 633–644, 2008.

[33] R. McKelvey. Intransitivities in multi-dimensional voting models and some implications for agenda control. *Journal of Economic Theory*, 12:472–482, 1976.

[34] Hervé Moulin, Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of Computational Social Choice*. Cambridge University Press, 2016.

[35] Roger B Myerson. *Game theory*. Harvard university press, 2013.

[36] Noam Nisan, Tim Roughgarden, Eva Tardos, and Vijay V. Vazirani. *Algorithmic Game Theory*. Cambridge University Press, New York, NY, USA, 2007.

[37] Matías Núñez and Jean-François Laslier. Bargaining through approval. *Journal of Mathematical Economics*, 60:63 – 73, 2015.

[38] Svetlana Obraztsova, Omer Lev, Evangelos Markakis, Zinovi Rabinovich, and Jeffrey S. Rosenschein. Distant truth: Bias under vote distortion costs. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, AAMAS '17, pages 885–892, Richland, SC, 2017. International Foundation for Autonomous Agents and Multiagent Systems.

[39] Svetlana Obraztsova, Evangelos Markakis, and David R. M. Thompson. Plurality voting with truth-biased agents. In Berthold Vöcking, editor, *Algorithmic Game Theory*, pages 26–37, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[40] Szilvia Pàpai. Strategyproof and nonbossy multiple assignments. *Journal of Public Economic Theory*, 3(3):257–271, 2001.

[41] Mark A. Satterthwaite and Hugo Sonnenschein. Strategy-proof allocation mechanisms at differentiable points. *The Review of Economic Studies*, 48(4):587–597, 1981.

[42] Mark Allen Satterthwaite. Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of economic theory*, 10(2):187–217, 1975.

A Proofs for the 1-Median Problem in $\mathbb{R}$ with Random-Breaking

Proof of Lemma 3.5. Trivially if $\bar{\pi}_1 = \pi_2$ then the best response for player 2 is $\bar{\pi}_1 + \sqrt{2}(\pi_2 - \bar{\pi}_1) = \pi_2$. Otherwise $\bar{\pi}_1 < \pi_2$. By scaling and translating we may assume $\bar{\pi}_1 = 0$ and $\pi_2 = 1$ and show that player 2's best response is $min\{u, \sqrt{2}\}$. It is straightforward to show that voter 2 should submit a location $x \geq 1$. The point Y is selected uniformly at random from $[0, x]$ and voter 2 has an expected cost of

$$E(|Y-1|) = P(Y \leq 1)E(1-Y|Y \leq 1) + P(Y \geq 1)E(Y-1|Y \geq 1) \tag{14}$$

$$= \frac{1}{x} \cdot \frac{1}{2} + \frac{x-1}{x} \cdot \frac{x-1}{2} \tag{15}$$

$$= \frac{x^2 - 2x + 2}{2x} \tag{16}$$

It is straightforward to verify that this function is uniquely minimized at $x = \sqrt{2}$ on the domain $[1, \infty]$. Therefore if $u \geq \sqrt{2}$ the unique best response for player 2 is $\sqrt{2}$. If $u < \sqrt{2}$ then it is easily verified that $E(|Y-1|)$ is strictly decreasing on the interval $[1, u]$ and therefore player 2's unique best response is $x = u$. Either way, player 2's unique best response is $min\{u, \sqrt{2}\}$. By the uniqueness of the best response, any more honest $\bar{\pi}_2$ yields a worse outcome for player 2 and therefore $min\{u, \sqrt{2}\}$ is player 2's unique minimally dishonest best response. □

Prior to proving Theorem 3.7 we need to establish a few properties for a minimally dishonest Nash equilibrium $\bar{\Pi}$.

Lemma A.1. *Suppose $|V| = 2$ and let $\bar{\Pi}$ be a minimally dishonest Nash equilibrium for the 1-Median problem when breaking ties uniformly at random given the sincere profile Π. If $\pi_1 \leq \pi_2$, then $\bar{\pi}_1 \leq \pi_2$ and $\pi_1 \leq \bar{\pi}_2$.*

Proof. For contradiction, suppose $\bar{\pi}_1 > \pi_2$. If $\bar{\pi}_2 \geq \pi_1$, then player 1 can obtain a strictly better outcome by being honest resulting in the point being selected uniformly at random from $[\pi_1, \bar{\pi}_2]$ instead of $[\bar{\pi}_2, \bar{\pi}_1]$ contradicting minimal dishonesty. Therefore $\bar{\pi}_2 < \pi_1$. If player 1 instead submits the honest π_1 then the point is selected uniformly at random from $[\bar{\pi}_2, \pi_1]$. Moreover, if $\pi_1 - \bar{\pi}_2 \leq \bar{\pi}_1 - \pi_1$ then player 1 likes this outcome at least as much as $r(\bar{\Pi})$. Therefore by minimal dishonesty

$$\pi_1 - \bar{\pi}_2 > \bar{\pi}_1 - \pi_1 = \bar{\pi}_1 - \pi_2 + \pi_2 - \pi_1. \tag{17}$$

Following from a symmetric argument for player 2,

$$\bar{\pi}_1 - \pi_2 > \pi_2 - \bar{\pi}_2 = \pi_2 - \pi_1 + \pi_1 - \bar{\pi}_2. \tag{18}$$

Combining both inequalities we obtain

$$\pi_1 - \bar{\pi}_2 > 2(\pi_2 - \pi_1) + \pi_1 - \bar{\pi}_2 > \pi_1 - \bar{\pi}_2 \tag{19}$$

a contradiction completing the proof of the lemma. □

Lemma A.1 is need to know how to apply Lemma 3.5. When combined with Lemma 3.5 we actually obtain the expected $\bar{\pi}_1 \leq \pi_1 \leq \pi_2 \leq \bar{\pi}_2$.

Lemma A.2. *Suppose $|V|$ is even and the point is selected with the 1-Median problem while breaking ties uniformly at random. Let $\bar{\Pi}$ be a minimally dishonest Nash equilibrium given the sincere profile Π . Then $a(\bar{\Pi}) \leq a(\Pi) \leq b(\Pi) \leq b(\bar{\Pi})$*

Proof. By symmetry, it suffices to show $b(\Pi) \leq b(\bar{\Pi})$. For contradiction suppose $b(\bar{\Pi}) < b(\Pi)$. Let v be a voter such that $\pi_v \geq b(\Pi)$. If $\bar{\pi}_v < b(\Pi)$ then $\bar{\pi}_v$ can obtain at least as good of an outcome by submitting $b(\Pi)$ contradicting minimal dishonesty. Therefore $\bar{\pi}_v \geq b(\Pi)$. This implies $|\{v : \bar{\pi}_v \geq b(\Pi)\}| \geq |\{v : \pi_v \geq b(\Pi)\}| \geq \frac{|V|}{2}$ and therefore $b(\bar{\Pi}) \geq b(\Pi)$, a contradiction completing the proof of the lemma. □

Lemma A.3. *Suppose $|V|$ is even and the point is selected with the 1-Median problem while breaking ties uniformly at random. Let $\bar{\Pi}$ be a minimally dishonest Nash equilibrium given the sincere profile Π . If $\pi_w \geq b(\Pi)$ then $\bar{\pi}_w \geq b(\bar{\Pi})$.*

Proof. For contradiction, suppose $\bar{\pi}_w < b(\bar{\Pi})$ implying $\bar{\pi}_w \leq a(\bar{\Pi})$. By Lemma A.2, $a(\bar{\Pi}) \leq a(\Pi) \leq \pi_w$. Voter w does not change the outcome if she submits $\bar{\pi}'_w = a(\bar{\Pi})$ and therefore by minimal dishonesty, $a(\bar{\Pi}) = \bar{\pi}_w < b(\bar{\Pi})$. Since $|\{v : \bar{\pi}_v \geq b(\bar{\Pi})\}| \geq \frac{|V|}{2}$, there must also be a voter y such that $\pi_y \leq a(\Pi)$ but $\bar{\pi}_w \geq b(\bar{\Pi})$. Through a symmetric argument $\bar{\pi}_y = b(\bar{\Pi})$.

Let $u = min_{v \in |V|: v \neq y}\{\bar{\pi}_v : \bar{\pi}_v \geq b(\bar{\Pi})\}$ and $l = max_{v \in |V|: v \neq w}\{\bar{\pi}_v : \bar{\pi}_y \leq a(\bar{\Pi})\}$. Finding the best responses $\bar{\pi}_w$ and $\bar{\pi}_y$ is now equivalent to the 2-player game with sincere preferences $\pi'_w = min\{u, \pi_w\}$ and $\pi'_w = max\{l, \pi_w\}$ where $\mathcal{X} = [l, u]$. By Lemmas 3.5 and A.1, $\bar{\pi}_y \leq \pi'_y \leq \pi'_w \leq \bar{\pi}_w$ contradicting that $\bar{\pi}_w = a(\bar{\Pi}) < b(\bar{\Pi}) = \bar{\pi}_y$ completing the proof of the lemma. □

Proof of Theorem 3.7. The lower bound is given by Theorem 3.6.

Let Π be a set of sincere preferences and $\bar{\Pi}$ be a corresponding minimally dishonest Nash equilibrium. Given a game with $|V|$ players we show how to reduce the game to 2 players that yields the same $r(\bar{\Pi})$. Finally we show that the price of deception of the 2-player game is at least as large as the price of deception of the $|V|$-player game. Therefore the price of deception is at most $\sqrt{2}$ by Theorem 3.6.

Let $|V| = 2n$ and index players so that $\pi_{-n} \leq \pi_{-n+1} \leq ... \leq \pi_{-1} \leq \pi_1 \leq ... \leq \pi_n$. Using Lemma A.3, it is straightforward to show that if $\pi_v \geq b(\Pi)$ then $\bar{\pi}_v = max\{b(\bar{\Pi}), \pi_v\}$. Therefore if $\bar{\Pi}$ is a minimally dishonest Nash equilibrium then $\bar{\pi}_{-n} \leq \bar{\pi}_{-n+1} \leq ... \leq \bar{\pi}_{-1} \leq \bar{\pi}_1 \leq ... \leq \bar{\pi}_n$ and players 1 and -1 are median voters with respect to Π and $\bar{\Pi}$. Construct the following 2-player game. Let $u = \bar{\pi}_2$ and $l = \bar{\pi}_{-2}$. Similar to Theorem 3.3, $u \geq b(\Pi) = \pi_1$

and $l \leq a(\Pi) = \pi_{-1}$. Players 1 and -1 have preferences $\pi'_v = \pi_v$. This is a valid game since $l \leq \pi'_{-1} \leq \pi'_1 \leq u$. It immediately follows that $\bar{\Pi}' = \{\bar{\pi}_{-1}, \bar{\pi}_1\}$ is a minimally dishonest Nash equilibrium for Π' with the same outcome as $\bar{\Pi}$ since otherwise $\bar{\Pi}$ is not a minimally dishonest Nash equilibrium for Π.

Let $C_v(\Pi, x) = ||\pi_v - x||_1$ be v's contribution to the social cost $C(\Pi, x)$ given location π_i and point x. Thus $E[C(\Pi, r(\bar{\Pi}))] = \sum_v E[C_v(\Pi, r(\bar{\Pi}))]$. Since $b(\Pi) = \pi_1 \leq \pi_v$ for $v \geq 1$,

$$C_v(\Pi, r(\Pi)) = C_1(\Pi, r(\Pi)) + \pi_v - \pi_1 \ \forall v \geq 1. \tag{20}$$

$$= UC_1(\Pi', r(\Pi')) + \pi_v - \pi_1 \ \forall v \geq 1. \tag{21}$$

By the triangle inequality,

$$E[C_v(\Pi, r(\bar{\Pi}))] \leq E[C_1(\Pi, r(\bar{\Pi}))] + \pi_v - \pi_1 \ \forall v \geq 1. \tag{22}$$

$$= E[C_1(\Pi', r(\bar{\Pi}'))] + \pi_v - \pi_1 \ \forall v \geq 1. \tag{23}$$

Symmetrically,

$$C_v(\Pi, r(\Pi)) = C_{-1}(\Pi', r(\Pi')) + \pi_{-1} - \pi_v \ \forall v \leq -1. \tag{24}$$

$$E[C_v(\Pi, r(\bar{\Pi}))] \leq E[C_{-1}(\Pi', r(\bar{\Pi}'))] + \pi_{-1} - \pi_v \ \forall v \leq -1. \tag{25}$$

We proceed by bounding the price of deception by grouping players v and $-v$. Observe that for $m_1, m_2, d_1, d_2 \geq 0$ that $\frac{m_1+m_2}{d_1+d_2} \leq max\{\frac{m_1}{d_1}, \frac{m_2}{d_2}\}$ where $\frac{0}{0} = 1$ and $\frac{x}{0} = \infty$ for $x > 0$. Therefore

$$\frac{E[C_v(\Pi, r(\bar{\Pi}))] + E[C_{-v}(\Pi, r(\bar{\Pi}))]}{C_v(\Pi, r(\Pi)) + C_{-v}(\Pi, r(\Pi))} \tag{26}$$

$$\leq \frac{E[C_1(\Pi', r(\bar{\Pi}'))] + E[C_{-1}(\Pi', r(\bar{\Pi}'))] + \pi_v - \pi_1 + \pi_{-1} - \pi_{-v}}{C_1(\Pi', r(\Pi'))] + C_{-1}(\Pi', r(\Pi'))] + \pi_v - \pi_1 + \pi_{-1} - \pi_{-v}} \tag{27}$$

$$\leq max\left\{\frac{E[C_1(\Pi', r(\bar{\Pi}'))] + E[C_{-1}(\Pi', r(\bar{\Pi}'))]}{C_1(\Pi', r(\Pi'))] + C_{-1}(\Pi', r(\Pi'))]}, \frac{\pi_v - \pi_1 + \pi_{-1} - \pi_{-v}}{\pi_v - \pi_1 + \pi_{-1} - \pi_{-v}}\right\} \tag{28}$$

$$\leq max\{\sqrt{2}, 1\} = \sqrt{2} \tag{29}$$

by Theorem 3.6 since Π' and $\bar{\Pi}'$ are from a 2-player game.

Thus the price of deception is

$$\frac{E[C(\Pi, r(\bar{\Pi}))]}{U(\Pi, r(\Pi))} = \frac{\sum_v E[C_v(\Pi, r(\bar{\Pi}))]}{\sum_v C_v(\Pi, r(\Pi))} \tag{30}$$

$$= \frac{\sum_{v\in[n]} \Big(E[C_v(\Pi, r(\bar{\Pi})) + E[C_{-v}(\Pi, r(\bar{\Pi}))]\Big)}{\sum_{v\in[n]} \Big(C_v(\Pi, r(\Pi)) + C_{-v}(\Pi, r(\Pi))\Big)} \tag{31}$$

$$\leq max_{v\in[n]} \frac{E[C_v(\Pi, r(\bar{\Pi})) + E[C_{-v}(\Pi, r(\bar{\Pi}))]}{C_v(\Pi, r(\Pi)) + C_{-v}(\Pi, r(\Pi))} \tag{32}$$

$$\leq \sqrt{2} \tag{33}$$

completing the proof of the theorem. □

B Proof of Theorem 4.6

The proof is similar identical to Theorem 4.6. Let $|V| = 2n$ for some integer $n \geq 2$. The sincere preferences are $\pi_v = (-1, 0)$ for odd indexed voters and let $\pi_v = (1, 0)$ for

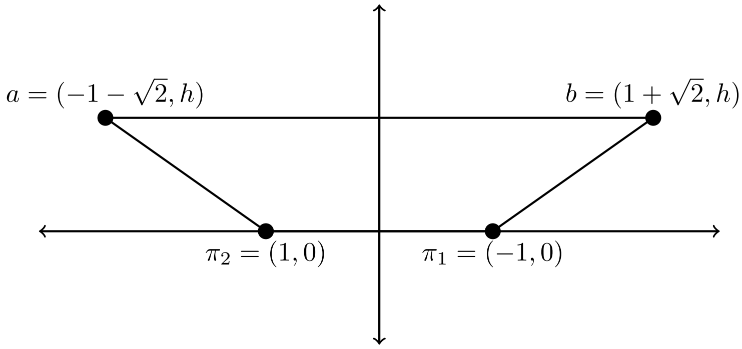

Figure 3: Preferences Showing the Price of Deception is ∞ for the 1-Median Problem with Random Tie-Breaking.

even indexed voters. Fix $h > 0$ and let $a = (-1-\sqrt{2}, h)$ and $b = (1+\sqrt{2}, h)$ and let $\mathcal{X} = conv.hull(\pi_1, a, b)$ as shown in Figure 3.

If everyone is honest then the selected point is $r(\Pi) = (0,0)$ with total cost $2|V|$. Suppose the odd indexed voters submit $\bar{\pi}_v = a$ and the even indexed voters submit $\bar{\pi}_v = b$. Then the point is selected uniformly at random from $[a, b]$ with total cost $2\sqrt{2}|V| + h \to \infty$ as $h \to \infty$. Therefore if $\bar{\Pi}$ is a minimally dishonest Nash equilibrium then the price of deception is ∞.

We now show that $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. By symmetry consider voter 2. As in Theorem 4.1 more than half the voters submit a height of h and therefore voter 2 cannot change the height of the point. If voter 2 submits $\bar{\pi}'_2 = (x, y)$ then the point will be selected uniformly at random between $(-1-\sqrt{2}, h)$ and (x, h). Let $X \sim Unif(-1, \sqrt{2}, x)$. Then voter 2's cost of the outcome is

$$c_2(\pi_2, r([\bar{\Pi}_{-2}, \bar{\pi}'_2])) = \sqrt[p_2]{E(|X-1|)^{p_2} + h^{p_2}} \tag{34}$$

which is minimized when $E(|X-1|)$ is minimized since h is a constant. Therefore the 2-dimensional game reduces to a 1-dimensional game. By Lemma 3.5, voter 2's minimally dishonest best response is to submit $x = 1+\sqrt{2}$ which requires $y = h$. Therefore voter 2 is submitting a minimally dishonest best response and $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. Thus the price of deception is ∞. □

C Proof of Theorem 4.8

We show the result for $p_v = 2$ for all v and explain how to generalize the result at the end. Consider the three points $\bar{\pi}_1 = (0,0), \bar{\pi}_2 = (\frac{3}{2sin(\alpha)}, \frac{3}{2cos(\alpha)})$, and $\bar{\pi}_3 = (-\frac{3}{2sin(\alpha)}, \frac{3}{2cos(\alpha)})$ where $\alpha < \frac{\pi}{2}$. Define $\mathcal{X} = conv.hull(\bar{\pi}_1, \bar{\pi}_2, \bar{\pi}_3)$. Suppose the sincere ideal points are given by $\pi_1 = (0,0), \pi_2 = (cos(\alpha), sin(\alpha))$, and $\pi_3 = (-cos(\alpha), sin(\alpha))$ as shown in Figure 4. With respect to the sincere data, the selected point is $r(\Pi) = (0, \frac{2sin(\alpha)}{3})$ with social cost

$$C(\Pi, r(\Pi)) = \sum_{i=1}^{3} ||r(\Pi) - \pi_i||_2^2 = \frac{2sin(\alpha)}{3} + 2\sqrt{1 - \frac{8sin^2(\alpha)}{9}} \leq 2 \tag{35}$$

Now consider the putative preferences given by $\bar{\Pi} = (\bar{\pi}_1, \bar{\pi}_2, \bar{\pi}_3)$. With respect to $\bar{\Pi}$ the selected point is $r(\bar{\Pi}) = (0, \frac{1}{cos(\alpha)})$. With respect to the sincere preferences Π, this has a

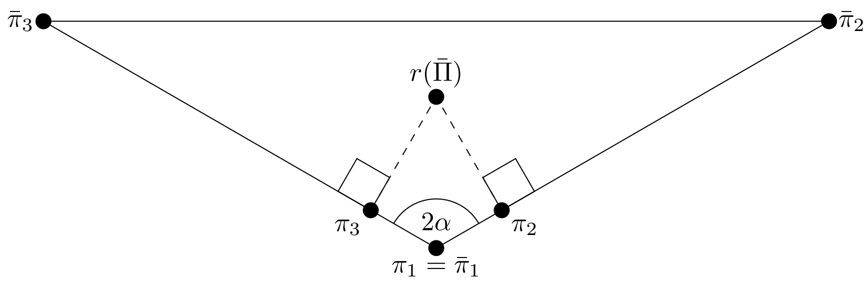

Figure 4: Preferences for Theorem 4.8

social cost of

$$C(\Pi, r(\bar{\Pi})) = \sum_{i=1}^{3} ||r(\bar{\Pi}) - \pi_i||_2^2 = \frac{1}{cos(\alpha)} + 2\sqrt{cos^2(\alpha) + \left(\frac{1}{cos(\alpha)} - sin(\alpha)\right)^2} \geq \frac{1}{cos(\alpha)}. \tag{36}$$

If $\bar{\Pi}$ corresponds to a minimally dishonest Nash equilibrium of Π, then the price of deception is at least

$$\frac{C(\Pi, r(\bar{\Pi}))}{C(\Pi, r(\Pi))} \geq \frac{1}{2cos(\alpha)} \rightarrow \infty \; as \; \alpha \rightarrow \frac{\pi}{2}. \tag{37}$$

It remains to show that $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. To do this, it suffices to show that any change to $\bar{\pi}_i$ yields a worse outcome for individual i. Start with $i = 1$. Suppose player 1 changes her preferences to $\bar{\pi}_1' = \bar{\pi}_1 + d$ where $||d|| \neq 0$. The location $\bar{\pi}_1'$ is in $\mathcal{X}$ and therefore $d_2 > 0$. After updating her preferences the point moves to $r([\bar{\Pi}_{-1}, \bar{\pi}_1']) = r(\bar{\Pi}) + \frac{1}{3}d$. This causes the point to move up and possibly to the left or the right. Regardless of p_1 this is worse for player 1 and therefore she is reporting a minimally dishonest best response.

The idea is similar to show players 2 and 3 are providing minimally dishonest best responses. By symmetry, it suffices to consider player 2. Suppose player 2 updates her preferences to $\bar{\pi}_2' = \bar{\pi}_2 - d$. Since $\bar{\pi}_2' \in \mathcal{X}$, $d \in cone\left((1,0),(\pi_2)\right)$. After updating her preferences, the selected point will be $r([\bar{\Pi}_{-2}, \bar{\pi}_2']) = r(\bar{\Pi}) - \frac{1}{3}d$. Let $B_2 = \{x : ||x - \pi_2||_2 \leq ||\pi_2 - r(\bar{\Pi})||_2$ be the set of points player 2 prefers to $r(\bar{\Pi})$. By construction, $\{x \in \mathbb{R} : r(\bar{\Pi}) - \frac{1}{3}dx\}$ is tangent to B_2 as shown in Figure 5 hence for all d where $||d|| > 0$, reporting $\bar{\pi}_2 - d$ would yield a worse solution for player 2. Thus she has given a minimally dishonest best response. This implies that $\bar{\Pi}$ is a minimally dishonest Nash equilibrium and therefore the price of deception converges to ∞ as α approaches $\frac{\pi}{2}$.

We now consider other $p_v \in (0, \infty)$. As observed previously, player 1 is minimally dishonest. In the construction given in Figure 4, $\bar{\pi}_2$ is placed such $A = \{x \in \mathbb{R} : r(\bar{\Pi}) - \frac{1}{3}dx\}$ is tangent to B_2 ensuring that voter 2 cannot alter her preferences to get a better outcome. However, if $p_2 \neq 2$ then the non-euclidean ball $B = \{x : ||\pi_2 - x||_{p_2} \leq ||\pi_2 - r(\bar{\Pi})||_{p_2}\}$ may overlap with A. However, if Figure 5 is stretched horizontally, then the set A converges to a horizontal line. Therefore for any finite norm, we can stretch the set of allowed feasible locations such that no individual can alter their preferences to get a better outcome. □

D The Partially Honesty Refinement

Another refinement in the voting literature if the partial honesty or truth-bias refinement.

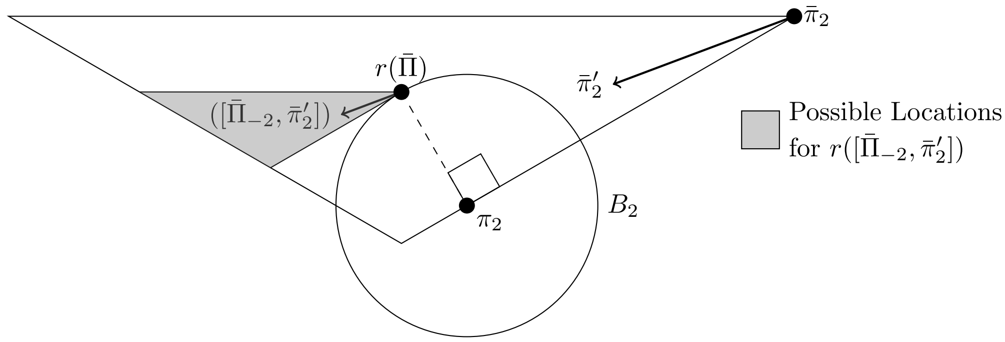

Figure 5: Possible Locations for $r([\bar{\Pi}_{-2}, \bar{\pi}'_2])$

Definition D.1. Let Π be the sincere preferences and let $\bar{\Pi}$ be a Nash equilibrium in the Strategic Spatial Social Choice Game. An individual v is partially dishonest if $c_v(\pi_v, r([\bar{\Pi}_{-v}, \pi_v])) > c_v(\pi_v, r(\bar{\Pi}))$.

A partially honest Nash equilibrium is a Nash equilibrium where all individuals are partially honest. By definition, a minimally dishonest Nash equilibrium is a partially honest Nash equilibrium. However, the reverse does not hold.

Theorem D.2. *There exist instances where a partially honest Nash equilibrium is not a minimally dishonest Nash equilibrium.*

Proof. Let $\mathcal{X} = [0,1]^2$, $|V| = 2k$ for some $k \geq 2$, and $\pi_v = (.5, 0)$ for all v and suppose $r(\cdot)$ is determined using the λ-1-Median problem with $\lambda = (.5, .5)$. By unanimity, $r(\Pi) = (.5, 0)$. From Theorem 4.2, it is straightforward to show that Π i s the only minimally dishonest Nash equilibrium. However, it is not the only partially honest Nash equilibrium.

Consider $\bar{\pi}_v = (1, 1)$ for half the voters and $\bar{\pi}_v = (0, 1)$ for the other half. With these preferences $r(\bar{\Pi}) = (.5, 1)$. Moreover, $\bar{\Pi}$ is a partially honest Nash equilibrium; If a voter v instead submits the honest $\pi_v = (0, 0)$ then the outcome moves to either $(0, 1)$ or $(1, 1)$. Both outcomes are worse for voter v and therefore v is partially honest completing the proof of the theorem. □

The partially honest equilibrium given in the proof of Theorem D.2 is unnatural. Every individual is lying about their π_{v2} coordinate despite there being no advantage to do so. This runs contrary to the experimental evidence discussed in 1.1 and thus we view partial honesty as ill-suited for examining the Strategic Spatial Social Choice Game.

Another approach to an honesty refinement is distorting utilities by penalizing individuals for being dishonest.

Definition D.3. Let $\epsilon > 0$. The ϵ-distorted cost of the point x with respect to sincere strategy π_v and submitted strategy $\bar{\pi}_v$ is $\bar{c}_v(\pi_v, x, \bar{\pi}_v) = c_v(\pi_v, x) + \epsilon||\bar{\pi}_v - \pi_v||_{p_v} = ||\pi_v - x||_{p_v} + \epsilon||\bar{\pi}_v - \pi_v||_{p_v}$.

It is straightforward to show that applying distorted costs results in minimally dishonest equilibria if players select their strategies from a finite set or more generally if $\{c_v(\pi_v, x) : x \in \mathcal{X}\}$ is finite and $||\bar{\pi}_v - \pi_v||$ is bounded. However in the setting of spatial social choice, using ϵ-distorted costs is not even a refinement on the set of Nash equilibria; applying these distorted costs can result in outcomes where individuals behave irrationally.

Theorem D.4. *Using ϵ-distorted costs in the Strategic Spatial Social Choice Game can result in irrational behavior.*

Proof. Let $\mathcal{X} = conv.hull\,((\pm 1, 1/\epsilon), (0,0))$, $V = \{1, 2, ..., 2k\}$ for some integer $k \geq 2$, and suppose $r(\cdot)$ is determined using the λ-1-Median problem with $\lambda = (.5, .5)$. Let $\pi_1 = (0,0)$, $\pi_v = (-1, 1/\epsilon)$ for $2 \leq v \leq k$, and $\pi_v = (1, 1/\epsilon)$ for $k+1 \leq v \leq 2k$ and let $p_v = 1$ for all v.

We begin by showing that Π is a Nash equilibrium when using ϵ-distorted costs. It is straightforward to verify that voter v is submitting a best response for all $v \geq 2$. If Voter 1 instead submits $\bar{\pi}_1 = (\bar{\pi}_{11}, \bar{\pi}_{22})$ then her ϵ-distorted cost is

$$||\pi_v - r([\bar{\pi}_1, \Pi_{-1}])||_{p_v} + \epsilon||\bar{\pi}_v - \pi_v||_{p_v} = \frac{\bar{\pi}_{11} - 1}{2} + \frac{1}{\epsilon} + \epsilon(\bar{\pi}_{11} + \bar{\pi}_{12}) \tag{38}$$

By definition of $\mathcal{X}$, $|\bar{\pi}_{11}| \leq \epsilon\bar{\pi}_{12}$. Therefore (38) is uniquely minimized by $\bar{\pi}_1 = \pi_1$ and Π is a Nash equilibrium when using distorted costs. However, Π is not a Nash equilibrium of the original game since voter 1 can obtain a strictly better outcome by submitting the unique best response $(1, 1/\epsilon)$. Therefore, rather than refining the set of equilibria, ϵ-distorted costs can result in irrational play. □

James P. Bailey
Singapore University of Technology and Design
Email: jamespbailey@gmail.com

Craig A. Tovey
Georgia Institute of Technology
Atlanta, Georgia
Email: craig.tovey@isye.gatech.edu