The Price of Anarchy in Spatial Social Choice

James P. Bailey
Rensselaer Polytechnic Institute
Troy, NY, United States
bailej6@rpi.edu

Craig A. Tovey
Georgia Institute of Technology
Atlanta, GA, United States
craig.tovey@isye.gatech.edu

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 use the *price of anarchy* to make finer distinctions than between "strategy-proof" and "manipulable." The price of anarchy measures how much strategic behavior can alter the cost in social choice. Supported by experimental economics data, our measure employs a novel *minimal dishonesty* criterion to refine the set of Nash equilibria. Using the price of anarchy, we study standard spatial selection rules and uncover a class of selection rules that are immune to the negative consequences of manipulation despite remaining manipulable. This is in contrast to standard approaches that sacrifice other beneficial properties, e.g., unbiased tie-breaking, to gain strategy-proofness. The concepts herein could be applied to other social choice scenarios in which a publicly known mechanism relies on private information.

KEYWORDS

social choice, facility location, manipulation, price of anarchy, Nash equilibrium refinements, mechanism design

ACM Reference Format:
James P. Bailey and Craig A. Tovey. 2025. The Price of Anarchy in Spatial Social Choice. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 9 pages.

1 INTRODUCTION

Like many social choice/voting mechanisms, spatial social choice is prone to manipulation. Typically such mechanisms are evaluated as either "manipulable" or "strategy-proof". In this paper, we use the price of anarchy to *quantify* the impact of manipulation and use it to identify social choice mechanisms that are manipulable yet still have powerful optimization guarantees with respect to the private, sincere preferences despite agents acting strategically. Thus, the price of anarchy offers finer distinctions between "manipulable" and "strategy-proof" and we propose that it be used as *one* of the criteria by which a selection rule is assessed.

In spatial social choice each agent has an ideal point in $\mathbb{R}^k$. The cost of a candidate to a voter is the Euclidean distance or some other metric distance between the two points. Spatial models are widely applied to facility location [19], preference aggregation [1, 11], political voting and policy selection [9, 12, 23, 28].

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

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 [10, 16, 17, 26]. However, there has not been much work toward understanding the quality of the outcomes obtained when the selection rule is subject to manipulation.

We use the price of anarchy [6, 13, 22] to measure how much strategic behavior impacts the quality of the outcome of a selection rule. 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 anarchy of the 1-Median problem would be at most 3.

We propose that the price of anarchy 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 total distance to the ideal points. If a point's cost has an intrinsic correspondence to its quality, then the price of anarchy has a correspondence to the capability of a selection rule to select a quality point. More generally, the price of anarchy measures a selection rule's ability to provide the expected sincere outcome despite strategic behavior. Like the computational complexity of manipulation [7, 29], the price of anarchy offers finer distinctions than simply between "manipulable" and "non-manipulable."

The concept of the price of anarchy 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 [30, 31] is irrelevant to the price of anarchy. 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 Our Contributions

In Section 3, we introduce the minimally dishonest Nash equilibrium refinement [2–5] to remove unnatural equilibria. The *minimal dishonesty* Nash equilibrium requires that an agent receives a strictly worse outcome if they are more honest; equivalently, an

agent only increases the size of their lie if it directly benefits them. Notably, the minimally dishonest Nash equilibrium is similar to another Nash refinement in social choice; when the set of candidates is finite, the minimally dishonest equilibrium is equivalent to partial honesty with ϵ-distorted costs [33, 34]. However, we show that in spatial social choice, the two concepts are distinct and the partial honesty with ϵ-distorted costs results in outcomes that are not equilibria before applying the refinement (Proposition 3.8). More importantly, we show that our minimally dishonest refinement removes spurious equilibria that are missed by partial honesty.

In Section 4, we use the price of anarchy and the minimally dishonest refinement to quantify the impact of manipulation for deterministic variants of the 1-Median problem. We show for arbitrary tie-breaking rules the price of anarchy for the 1-Median problem is infinity (Theorem 4.2); this indicates that manipulation completely erodes any guarantees for the 1-median problem and that arbitrarily poor outcomes can be obtained. We show that introducing a "left" or "right" bias into the tie-breaking rule causes the 1-Median problem to become strategy-proof (Theorem 4.3). However, introducing bias into a mechanism is undesirable, and instead, we use the price of anarchy to design a mechanism that maintains optimality guarantees despite being manipulable. Specifically, we show that limiting allowed outcomes to a hyperrectangle causes the price of anarchy to decrease to 1 in the 1-Median problem (Theorem 4.6). This means that, while the mechanism itself is manipulable, manipulation has no impact on the quality of the outcome; the social cost will remain the same with respect to the private, sincere preferences even when agents are strategic and misrepresent their preferences. To the best of our knowledge, this is the first time a manipulable social choice algorithm has been shown to guarantee optimal solutions with respect to sincere data. We also remark that submitting preferences within a hyperrectangle is natural as, for example, social policies do not limit your rating of economic policies.

In Section 5, we use the price of anarchy to study the 1-mean problem. Like the 1-median problem, the price of anarchy is infinity in general, and arbitrarily poor outcomes can be obtained when individuals are strategic (Theorem 5.1). Also like the 1-median problem, we show that better mechanism design can lessen the impact of manipulations; we show that if the set of outcomes is limited to a hyperrectangle then the price of anarchy decreases to $O(|V|)$ where V is the set of voters/agents; the increase in social cost due to manipulation is bounded by a linear factor.

Finally, in Section 6, we study the spatial social choice problem when the outcome is found by minimizing the l_2-norm between agents' ideal points and the outcome. In this setting, we show that the price of anarchy is always infinity, even when the set of outcomes is limited to a hyperrectangle (Theorem 6.1).

Importantly, our results demonstrate that not all forms of manipulation are equal; if a mechanism designer isn't careful then the 1-median, 1-mean, and l_2 selection rules can result in arbitrarily poor outcomes. However, by using the price of anarchy, our proposed methods allow us to analyze/design mechanisms where manipulation has a small, or even no, impact on the quality of the outcome obtained. This is especially important in social choice/voting mechanisms where many impossibility results exist that require certain desirable properties to be sacrificed to achieve strategy-proofness.

2 NOTATION

An instance of the Spatial Social Choice problem consists of a compact, convex domain $\mathcal{X} \subseteq \mathbb{R}^k$ of feasible outcomes and a set V of agents. Each agent $v \in V$ has an ideal point $\pi_v \in \mathcal{X}$ representing v's most preferred point in $\mathcal{X}$. We refer to $\Pi = \{\pi_v\}_{v \in V} \in \mathcal{X}^{|V|}$ as the sincere *preference profile*. Agents then submit their preferences to a publicly known selection rule r. If the selection rule is deterministic, then the outcome is a single outcome $r(\Pi) \in \mathcal{X}$.

We consider three common selection rules along with possible tie-breaking rules. Each selection rule is a minimizer of a social cost function $C : \mathcal{X}^{|V|} \times \mathcal{X} \to \mathbb{R}$, e.g., if C has a unique optimizer, then $r(\Pi) := \text{argmin}_{x \in \mathcal{X}} C(\Pi, x)$. As a result, these selection rules are often associated with the *social cost* of an outcome; each selection rule we consider aims to minimize some total *distance* from agents' ideal points to the selected outcome. The three primary selection rules we consider are defined by the following social cost functions.

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

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

$$C(\Pi, x) := \sum_{v \in V} ||\pi_v - x||_2 \quad (l_2\text{-Norm})$$

Only the 1-Mean problem is guaranteed a unique optimizer. For the other two problems we consider deterministic tie-breaking rules which we introduce in their respective sections.

We consider the most common spatial choice problem where an agent prefers outcomes closer to their ideal point. In this paper, we assume agent v evaluates the distance between an outcome and their ideal point using the l_{p_v}-*norm*[1] for some $p_v \in (0, \infty)$. Formally, for agent v, we assume that their cost for the outcome $x \in \mathcal{X}$ is $c_v(x) := ||\pi_v - x||_{p_v} = (\sum_i (\pi_{vi} - x_i)^{p_v})^{1/p_v}$ for some $p_v \in (0, \infty)$. In standard analyses, p_v is equal to 2 corresponding to the Euclidean norm. Our results hold for any $p_v \in (0, \infty)$ and even when agents use different p-norms, e.g., if $p_v \neq p_{v'}$ for some $v' \neq v$.

2.1 Spatial Social Choice Game

In practice, agents will not necessarily submit honest preferences. For example, consider $\mathcal{X} = [0, 1]$ and $|V| = 2$ with sincere ideal points $\pi_1 = 0$ and $\pi_2 = 0.5$. Suppose further that we are solving the 1-Mean problem. If agents are sincere then the (truthful) outcome is $r(\Pi) = r((0, 0.5)) = \text{argmin}_{x \in \mathcal{X}} \sum_{v \in V} ||\pi_v - x||_2^2 = 0.25$, which is $c_2(r(\Pi)) = ||\pi_2 - r(\Pi)||_{p_2} = 0.25$ away from agent 2's ideal point.

However, agent 2 can misrepresent their ideal point to result in a strictly better outcome for agent 2; if agent 2 instead submits the submitted ideal point $\bar{\pi}_2 = 1$ resulting in the submitted preference profile $\bar{\Pi} = (0, 1)$, then the (submitted/actual) outcome is $r(\bar{\Pi}) = 0.5$, which is $c_2(r(\bar{\Pi})) = 0$ away from agent 2's true ideal point. This manipulation of submitted preferences results in the following Strategic Spatial Social Choice Game.

Strategic Spatial Social Choice Game

[1] p_v corresponds to a norm only if $p_v \geq 1$ since the triangle inequality is violated when $p_v < 1$. However, the triangle inequality is not used in our proofs and we allow $p_v \in (0, \infty)$.

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

Games are typically understood via their Nash equilibria. A set of submitted preferences $\bar{\Pi}$ forms a pure strategy Nash equilibrium if no agent 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

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

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

We remark that the study of games often considers mixed equilibria – they allow agents to play distributions over the set of pure strategies. In games where the number of pure strategies is finite, mixed equilibria are useful to guarantee the existence of Nash equilibria. In our setting, since the strategy space is compact and convex, and since the selection rule r is a continuous function of agent preferences for the three selection rules we consider, it is straightforward to show a pure Nash equilibrium exists in the Strategic Social Choice Game using Brouwer's or Kakutani's fixed point theorems – in fact, in Proposition 3.1 we demonstrate that the set of Nash equilibria is dense. Since it is well-known that pure strategies are always a best response in game theory, and since pure strategy equilibria exist, we only consider pure strategies. This is also consistent with prior literature studying spatial social choice (see e.g., [16, 17, 33]).

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 $r(\Pi) = \sum_{v \in V} \frac{\pi_v}{|V|}$. *Consider the domain* $\mathcal{X} = \{x \in \mathbb{R}^2 : (0,0) \leq x \leq (1,1)\}$ *and the sincere ideal points* $\pi_1 = (0,0), \pi_2 = (0, \frac{1}{3})$, *and* $\pi_3 = (\frac{1}{3}, 0)$. *With respect to these preferences, the selected outcome is* $r(\Pi) = (\frac{1}{9}, \frac{1}{9})$. *This corresponds to a total 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.*

In this example, agent 1 would like to move the sincere outcome $r(\Pi)$ *to the lower left, agent 2 to the upper left, and agent 3 to the lower right. A Nash equilibrium where agents attempt to do 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 outcome is* $r(\bar{\Pi}) = (\frac{1}{3}, \frac{1}{3})$. *With respective to the submitted preferences, it appears that the social cost is* $C(\bar{\Pi}, r(\bar{\Pi})) = \sum_{i=1}^{3} ||\bar{\pi}_i - r(\bar{\Pi})||_2^2 = \frac{4}{3}$. *However, with respect to the sincere preferences* Π, *the outcome actually costs* $C(\Pi, r(\bar{\Pi})) = \sum_{i=1}^{3} ||\pi_i - r(\bar{\Pi})||_2^2 = \frac{4}{9}$.

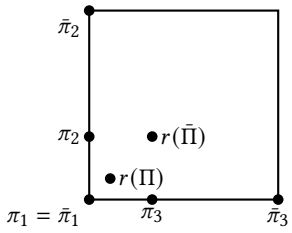

Figure 1: Sincere and submitted preferences for Example 2.1.

To see that $\bar{\Pi}$ *corresponds to a Nash equilibrium, consider agent 1. If agent 1 alters their submitted ideal point* $\bar{\pi}_1$, *then they must move it up or to the right. Such an action causes the point to move up or to the right respectively, which is farther away from* π_1 *with respect to every* l_p*-norm, and therefore agent* 1 *cannot alter their submitted preferences to get a better result. Agents 2 and 3 provide best responses by similar reasoning and* $\bar{\Pi}$ *is a pure strategy Nash equilibrium.*

2.2 Price of Anarchy

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. In general, we aim to bound how much worse an outcome obtained from a Nash equilibrium is relative to the sincere outcome using the *price of anarchy*. The price of anarchy is a worst-case analysis of how much manipulation can impact the social cost and provides an indicator of a selection's rule ability to provide the promised outcome.

Definition 2.2. Let r be a selection rule that minimizes the real-valued cost function C over the domain $\mathcal{X}$ and let V be a set of agents. Let $NE(\Pi)$ denote the set of equilibria of the Strategic Spatial Social Choice Game with r given the sincere profile Π. Then the price of anarchy of the selection rule r is

$$\sup_{\Pi \in \mathcal{X}^{|V|}} \sup_{\bar{\Pi} \in NE(\Pi)} \frac{C(\Pi, r(\bar{\Pi}))}{C(\Pi, r(\Pi))} \qquad \text{(The Price of Anarchy)}$$

Proving that the price of anarchy is u requires two parts. First, there must be an instance (or family of instances) showing that manipulation can cause the social cost to increase by a factor of u (or arbitrarily close to u) indicating the price of anarchy is at least u. Second, we must show that there cannot be an instance where manipulation can cause social cost to increase by a factor more than u. Example 2.1 demonstrated that the 1-Mean problem with 3 agents may result in the social cost increasing by a factor of 3 and therefore the price of anarchy is at least 3. To show a price of anarchy of 3 we would also have to provide a proof that manipulation cannot cause the social cost to increase by a factor of more than 3.

3 THE MINIMALLY DISHONEST EQUILIBRIUM

Prior to proving the price of anarchy for various spatial social choice mechanisms, we first demonstrate a need to refine the set of Nash equilibria. For spatial social choice, we opt to use the minimal dishonesty refinement [2–5]. We also briefly discuss other common

Nash equilibria refinements in social choice and demonstrate that they are ill-suited for the spatial social choice problem.

The study of strategic behavior in social choice is typically limited to determining whether a mechanism is "strategy-proof" or "manipulable" – i.e., whether the sincere profile Π is a Nash equilibrium – and little research has focused on understanding non-truthful equilibria. We believe this to be, at least partially, because the Nash equilibrium solution concept, at least as we have described it, often makes little sense in social choice games. For instance, for the 1-Median problem every $x \in \mathcal{X}$ is obtainable by at least one Nash equilibrium. Hence the standard Nash equilibrium has no predictive power in this setting.

Proposition 3.1. *Consider the* 1*-Median problem on the domain* $\mathcal{X}$ *with* $|V| \geq 3$ *agents. For every sincere preference profile* $\Pi \in \mathcal{X}^{|V|}$*, the submitted preference profile* $\bar{\Pi} = (x, x, ..., x)$ *is a Nash equilibrium with outcome* $r(\bar{\Pi}) = x$ *for all* $x \in \mathcal{X}$*.*

The proof follows immediately since $|V| \geq 3$ implies x is a median even after a single voter moves. Proposition 3.1 demonstrates that many outcomes of the Strategic Spatial Social Choice Game make little sense. Indeed, voting mechanisms, including mechanisms outside of spatial social choice, that have a near-unanimity property – e.g., if all but one agent reach a consensus then the outcome will be determined by that consensus – will have similar meaningless equilibria. Further, this near-unanimity property is common in most voting mechanisms, e.g., every Condorcet-consistent[2] voting mechanism, a common property in social choice, with $n \geq 3$ agents has this property. Therefore, meaningless Nash equilibria are quite common in social choice games.

We believe these absurd equilibria are the biggest obstacle to understanding the effects of strategic behavior.

To eliminate these spurious equilibria, we introduce the *Minimally Dishonest* Nash equilibrium refinement. Informally, a minimally dishonest agent v would never submit $\bar{\pi}_v$ if there exists a more honest $\bar{\pi}'_v$ that yields at least as good of an outcome for v. We believe this refinement is intuitive. More importantly, the notion of a refinement introducing a preference for honesty has a precedent in the voting literature [14, 15, 21, 24, 27, 32–34] and is backed by a large amount of experimental evidence (e.g., [8, 18, 20, 25]). 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 3.2. Let Π be a sincere preference profile and let $\bar{\Pi}$ be a submitted preference profile in the Strategic Spatial Social Choice Game. An agent v is minimally dishonest with respect to $\bar{\Pi}$ if $||\pi_v - \bar{\pi}'_v||_{p_v} < ||\pi_v - \bar{\pi}_v||_{p_v}$ implies $c_v(r([\bar{\Pi}_{-v}, \bar{\pi}'_v])) > c_v(r(\bar{\Pi}))$, i.e., if submitting the more honest $\bar{\pi}'_v$ would result in a worse outcome.

An agent is minimally dishonest if being more honest always results in a strictly worse outcome for the agent. Equivalently, an agent will only increase the size of their lie if it strictly benefits them. A minimally dishonest Nash equilibrium is a Nash equilibrium where every agent is minimally dishonest. We remark that our Nash equilibrium refinement is consistent with the literature's focus on *strategy-proof* mechanisms.

[2]A mechanism is Condorcet-consistent if the existence of an $x \in \mathcal{X}$ where for every $x' \neq x$ a majority of agents prefer x to x' implies the mechanism selects x.

Definition 3.3. 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.

Interestingly, strategy-proof mechanisms frequently have the same type of absurd equilibria as demonstrated in Proposition 3.1. Despite the lack of unique equilibria, researchers typically assume all agents will be honest in strategy-proof mechanisms. Our minimally dishonest equilibrium refinements strongly support this assumption; we show that a mechanism is strategy-proof if and only if honesty is the unique minimally dishonest Nash equilibrium.

Proposition 3.4. *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 a best response for agent v to submit the honest π_v regardless of all other preferences. Therefore π_v is the unique minimally dishonest best response for agent 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. □

We remark that there are other refinements of Nash equilibria in the literature on social choice. The primary refinement used is the partial honesty or equivalently the truth-bias refinement. In this refinement, truthfulness is evaluated in a binary sense; individual v would not submit $\bar{\pi}_v$ if submitting the sincere ideal point π_v would yield at least as good of an outcome. Notably, this refinement does not distinguish between large and small lies.

Definition 3.5. Let Π be a sincere preference profile and let $\bar{\Pi}$ be a submitted preference profile in the Strategic Spatial Social Choice Game. An individual v is partially honest with respect to $\bar{\Pi}$ if $c_v(r([\bar{\Pi}_{-v}, \pi_v)]) > c_v(r(\bar{\Pi}))$ when $\bar{\pi}_v \neq \pi_v$, i.e., if submitting the honest π_v would result in a worse outcome.

A partially honest equilibrium is an equilibrium where every individual is partially honest. Trivially, every minimally dishonest Nash equilibrium is a partially honest equilibrium. However, in Proposition 3.6, we demonstrate that not every partially honest equilibrium is a minimally dishonest equilibrium, implying that the minimally dishonest refinement provides a smaller set of equilibria while remaining consistent with literature on social aversion to lying. Perhaps more importantly, the partially honest equilibrium constructed in Proposition 3.6 is still quite unnatural.

Proposition 3.6. *There exist instances where a partially honest Nash equilibrium is not a minimally dishonest Nash equilibrium.*

Proof. We consider the 1-Median problem with agents $V = \{1, ..., 2k\}$. It is well-known that the 1-Median problem may have ties when there are an even number of agents and the set of optimal outcomes with respect to the profile Π is the hyperrectangle $\{x : a(\Pi) \leq x \leq b(\Pi)\}$ where $a_i(\Pi) = \text{argmin}\{\pi_{vi} : |\{v' : \pi_{v'i} \leq \pi_{vi}\}| \geq |V|/2\}$ and $b_i(\Pi) = \text{argmax}\{\pi_{vi} : |\{v' : \pi_{v'i} \geq \pi_{vi}\}| \geq |V|/2\}$, i.e., $\{x : a(\Pi) \leq x \leq b(\Pi)\}$ is the set of medians with respect to Π. For this proposition, we break ties by selecting the midpoint $r(\Pi) = 0.5 \cdot a(\Pi) + 0.5 \cdot b(\Pi)$.

Consider the sincere profile $\pi_v = (0.5, 0)$ for all $v \in V$. The only median with respect to Π is $r(\Pi) = (0.5, 0)$. Next, consider the submitted profile $\bar{\Pi}$ where $\bar{\pi}_v = (0, 1)$ for $v = 1, ..., k$ and where $\bar{\pi}_v = (1, 1)$ for $v = k+1, ..., 2k$. The set of medians is $\{x : (0, 1) \leq x \leq (1, 1)\}$ resulting in the outcome $r(\bar{\Pi}) = (0.5, 1)$.

We now show that the submitted profile $\bar{\Pi}$ is a partially honest Nash equilibrium. By symmetry, it suffices to show that agent $v = 2k$ is providing a partially honest best response. If agent $2k$ instead submits $\bar{\pi}' = ([\bar{\pi}']_1, [\bar{\pi}']_2)$, then the new set of medians will be given by $\{x : (0, 0) \leq x \leq ([\bar{\pi}']_1, 1)\}$ yielding the outcome $([\bar{\pi}'_v]_1/2, 1)$. This outcome is no better for agent $2k$ and therefore $\bar{\Pi}$ corresponds to a Nash equilibrium. Further, agent $2k$ is partially honest since submitting the honest $(0.5, 0)$ would result in the outcome $(0.25, 1)$ – an outcome that is worse for agent $2k$ with respect to every p-norm. However, $\bar{\Pi}$ is not a minimally dishonest Nash equilibrium since agent $2k$ can submit the more honest $\bar{\pi}' = (1, 0)$ and obtain the same outcome. □

The partially honest equilibrium in Proposition 3.6 is quite unrealistic. Despite all agents agreeing on the ideal outcome, there exist partially honest equilibria where the final outcome is quite far from the actual preferences. The minimally dishonest equilibrium concept removes these unrealistic equilibria. Moreover, in Theorem 4.6, we show this selection rule has a price of anarchy of 1, implying $(0.5, 0)$ is the only possible minimally dishonest outcome for the preferences in Proposition 3.6. Specifically for the equilibrium in Proposition 3.6, no agent is minimally dishonest since each agent can be more honest by changing the second coordinate of their submitted ideal point to 0; if each agent does so, the resulting profile is minimally dishonest with $r(\bar{\Pi}) = (0.5, 0) = r(\Pi)$.

Similar absurd partially honest equilibria were observed in classical voting settings in [33], and a different adjustment was introduced. The idea proposed in [33] is to distort the utility/cost by penalizing individuals for being dishonest.

Definition 3.7. Let $\epsilon > 0$. The ϵ-distorted cost of the outcome $x \in \mathcal{X}$ with respect to sincere ideal point π_v and submitted ideal point $\bar{\pi}_v$ is $\bar{c}_v(x, \bar{\pi}_v) = c_v(x) + \epsilon \cdot ||\bar{\pi}_v - \pi_v||_{p_v} = ||\pi_v - x||_{p_v} + \epsilon \cdot ||\bar{\pi}_v - \pi_v||_{p_v}$.

It is straightforward to show that applying distorted costs with sufficiently small ϵ results in minimally dishonest equilibria if individuals select their strategies from a finite set, which is the case in candidate-based voting [4, 33] and stable matching mechanisms [5]. However, in the setting of spatial social choice, using ϵ-distorted costs is not a proper refinement on the set of Nash equilibria.

Proposition 3.8. *The Nash equilibria obtained after applying ϵ-distorted costs to the Strategic Spatial Social Choice Game are not necessarily Nash equilibria.*

Proof. We use the same selection rule as in Proposition 3.6. Let $\mathcal{X}$ be the triangle formed by the convex-hull of $((\pm 1, 1/\epsilon), (0, 0))$, and let $V = \{1, 2, ..., 2k\}$ for some integer $k \geq 2$. Suppose $\pi_1 = (0, 0)$, $\pi_v = (-1, 1/\epsilon)$ for $2 \leq v \leq k$, and $\pi_v = (1, 1/\epsilon)$ for $k+1 \leq v \leq 2k$ yielding $r(\Pi) = (0.5, 1/\epsilon)$ as depicted in Figure 2. Observe that Π is not a Nash equilibrium; agent 1 can obtain a strictly better outcome by submitting the unique best response $\bar{\pi}_1 = (1, 1/\epsilon)$ – regardless of which p-norm agent 1 is using – resulting in the outcome $r([\Pi_{-1}, \bar{\pi}_1]) = (0, 1/\epsilon)$.

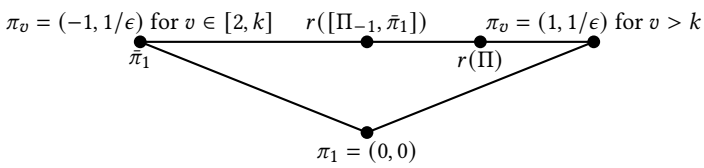

Figure 2: Preferences for Proposition 3.8.

However, Π is a Nash equilibrium when using ϵ-distorted costs if agents evaluate their costs with respect to the l_1 norm ($p_v = 1$ for all v): It is straightforward to verify that agent v is submitting a best ϵ-distorted response for all $v \geq 2$; they are submitting a best response and their distorted costs are 0. If agent 1 instead submits $\bar{\pi}_1 = (\bar{\pi}_{11}, \bar{\pi}_{22})$ then the new outcome is $r([\Pi_{-1}, \bar{\pi}_1]) = ((1 + \bar{\pi}_{11})/2, 1/\epsilon)$. As a result, agent 1's ϵ-distorted cost is

$$||\pi_1 - r([\Pi_{-1}, \bar{\pi}_1])||_1 + \epsilon \cdot ||\bar{\pi}_v - \pi_v||_1 \quad (1)$$

$$= \frac{1 + \bar{\pi}_{11}}{2} + \frac{1}{\epsilon} + \epsilon \cdot (|\bar{\pi}_{11}| + \bar{\pi}_{12}) \quad (2)$$

At $\bar{\pi}_1 = \pi_1 = (0, 0)$, this evaluates to $1/2 + 1/\epsilon$. For $\bar{\pi}_{11} > 0$, the ϵ-distorted cost is strictly more than $1/2 + 1/\epsilon$ since $\bar{\pi}_{22} \geq 0$. For $\bar{\pi}_{11} < 0$, we have that $|\bar{\pi}_{11}| = -\bar{\pi}_{11}$ and $\bar{\pi}_{12} \geq -\bar{\pi}_{11}/\epsilon$ by construction of $\mathcal{X}$. For this case, the ϵ-distorted cost is

$$\frac{1 + \bar{\pi}_{11}}{2} + \frac{1}{\epsilon} - \epsilon \cdot \bar{\pi}_{11} + \epsilon \cdot \bar{\pi}_{12}$$
$$\geq 1/2 + 1/\epsilon - (1/2 + \epsilon) \cdot \bar{\pi}_{11} > 1/2 + 1/\epsilon$$

since $\bar{\pi}_{11} < 0$. Therefore (1) is uniquely minimized by $\bar{\pi}_1 = \pi_1 = (0, 0)$ and Π is a Nash equilibrium when using distorted costs despite Π not being a Nash equilibrium. This means that the ϵ-distorted costs concept is not a Nash refinement. □

By Propositions 3.6 and 3.8, the partial-honesty refinement and the distorted costs are ill-suited for the Strategic Social Choice Game. Using the Minimally Dishonesty refinement, we study the impact of manipulation on classical spatial social choice mechanisms.

4 1-MEDIAN PROBLEM

The 1-Median problem minimizes the social cost $C(\Pi, x) = \sum_{v \in V} ||\pi_v - x||_1$. The outcome will necessarily be a median for the set of submitted ideal points. As in Propositions 3.6 and 3.8, let $a_i(\Pi) = \text{argmin}\{\pi_{vi} : |\{v' : \pi_{v'i} \leq \pi_{vi}\}| \geq |V|/2\}$ and $b_i(\Pi) = \text{argmax}\{\pi_{vi} : |\{v' : \pi_{v'i} \geq \pi_{vi}\}| \geq |V|/2\}$. Then every $x \in [a(\Pi), b(\Pi)]$ minimizes $C(\Pi, x)$. When $|V|$ is odd, $a(\Pi) = b(\Pi)$ yielding a unique solution. However, when $|V|$ is even, multiple optimal solutions may exist. We consider deterministic tie-breaking rules.

Definition 4.1. 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 the outcome $r_i(\Pi) = (1 - \lambda_i) \cdot a_i(\Pi) + \lambda_i \cdot b_i(\Pi)$ for $i = 1, ..., k$.

We first show that the λ-1-median selection rule can result in a price of anarchy of ∞ when implemented poorly, i.e., some variants of the λ-1-median problem are incapable of guaranteeing any meaningful outcome when individuals are strategic.

Theorem 4.2. *Suppose $\lambda \in (0,1)^k$ and $|V| \geq 4$ is even. When individuals are minimally dishonest, the price of anarchy of the λ-1-Median problem in $\mathbb{R}^k$ is ∞ for $k \geq 2$. Specifically, for all $\epsilon \in (0,1]$ there exists an instance with a price of anarchy of $\frac{|V|-\epsilon}{\epsilon} \to \infty$ as $\epsilon \to 0$.*

Proof. It suffices to show the result for $k = 2$ since the lower dimensional example can always be embedded in higher dimensions.

Given the tie-breaking rule $\lambda = (\lambda_1, \lambda_2)$, let $a = (\frac{\lambda_1 - 1}{\lambda_1}, 1)$ and $b = (1,1)$. Let $\mathcal{X} = conv.hull(a, b, \vec{0})$ where $conv.hull(\cdot)$ denotes the convex hull. Let $|V| = 2 \cdot n$ for some integer $n \geq 2$. Suppose an arbitrary voter has the preference $\pi_v = (0, \epsilon)$ for some $\epsilon \in (0,1]$ and all other voters have the preference $\pi_v = (0,0)$ yielding the outcome $r(\Pi) = (0,0)$ with a social cost of ϵ. We now describe a minimally dishonest Nash equilibrium with respect to Π that has a social cost of $|V| - \epsilon$. Consider the submitted preferences given by $\bar{\pi}_v = a = (\frac{\lambda_1 - 1}{\lambda}, 1)$ for $v = 1, ..., n$ and $\bar{\pi}_v = b = (1,1)$ for $v = n+1, ..., 2n$.

With respect to $\bar{\Pi}$, the medians are $\left\{x : \left(\frac{\lambda_1 - 1}{\lambda_1}, 1\right) \leq x \leq (1,1)\right\}$ and $r(\bar{\Pi}) = (0,1)$ since $r_1(\bar{\Pi}) = \lambda_1 \cdot a_1 + (1 - \lambda_1) \cdot b_1 = 0$. This outcome has a sincere cost of $|V| - \epsilon$.

We now show that $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. Consider agent v that submits $\bar{\pi}_v = b$. Regardless of how they alter their submitted information, more than half of the agents submit a height of 1 and therefore they cannot change the height of the outcome. Moreover, if they alter their preferences at all then the outcome moves to the left corresponding to a worse outcome for her. Formally; if they submit $\bar{\pi}'_v$, then the set of medians changes to $\left\{x : \left(\frac{\lambda_1 - 1}{\lambda_1}, 1\right) \leq x \leq (\bar{\pi}'_{v1}, 1)\right\}$, which results in the worse outcome $r([\bar{\Pi}_{-v}, \bar{\pi}'_v] = (\lambda_1 - 1 + (1 - \lambda_1) \cdot \bar{\pi}'_{v1}, 1)$ when $\bar{\pi}'_v \neq \bar{\pi}_v$. Therefore v is providing a minimally dishonest best response. Symmetrically, an agent v submitting $\bar{\pi}_v = a$ is providing a minimally dishonest best response and $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. □

The result of Theorem 4.2 is disappointing; it implies that, in general, the 1-median selection rule is incapable of guaranteeing good outcomes. However, we show that this issue can be resolved by introducing a "left" or "right" bias into the tie-breaking rule, e.g., by always selecting the optimal solution furthest to the left.

Theorem 4.3. *Suppose $\mathcal{X} = \{x : a \leq x \leq b\}$ is a hyperrectangle and $\lambda \in \{0,1\}^k$. Then the λ-1-median problem is strategy-proof.*

Proof. Since $\mathcal{X}$ is a hyperrectangle, the decision process for each agent is separable with respect to each axis, i.e., $\bar{\pi}_{vi}$ is independent of $\bar{\pi}_{vj}$ for $j \neq i$. Further, the p_v norm is such that decreasing the distance with respect to a single coordinate (while keeping all other coordinates the same) causes the total distance to decrease. Finally, the selection rule is separable, and therefore the strategic social choice game is separable. As a result, it suffices to show the result when $k = 1$, i.e., when all ideal points are on the line segment $[a, b]$. We now show that π_v is a best response for every agent v. Without loss of generality, assume that $\lambda = 1$ implying $r(\Pi) = b(\Pi)$.

First, consider an agent v where $\pi_v = b(\Pi)$. This agent is receiving their preferred outcome and π_v is a best response.

Next, consider an agent v where $\pi_v > b(\Pi)$. If agent v shifts their ideal point to the left, then the set of median points moves to the left or remains unchanged. Neither outcome is better for agent v. If instead agent v shifts their ideal point to the right, the outcome does not change since v is not a median voter. Thus, v is providing a best response when honest.

Finally, consider an agent v where $\pi_v < b(\Pi)$. If agent v shifts their ideal point to the right, then the set of median points moves to the right or remains unchanged. Neither outcome is better for agent v. Further, if agent v shifts their ideal point to the left, then $b(\Pi)$ remains a median. While this shift may cause $a(\Pi)$ to decrease, the outcome will not change since $r(\bar{\Pi}) = b(\bar{\Pi})$. Therefore π_v is also a best response for agent v. Thus, honesty is a best response for every agent. □

Corollary 4.4. *Suppose $\lambda \in \{0,1\}^k$. Then the λ-1-median problem is strategy-proof.*

Proof. Let $a, b \in \mathbb{R}^k$ be such that $\mathcal{X} \in [a, b]$. By Theorem 4.3, Π is a Nash equilibrium with respect to Π in the domain $[a, b]$, i.e., π_v is a best response for each agent v. This implies π_v is also a best response in the domain $\mathcal{X}$ since $\mathcal{X} \subseteq [a, b]$ and since $\pi_v \in \mathcal{X}$. □

Further, when $|V|$ is odd, there is a unique median and every tie-breaking rule is equivalent, implying that the 1-median problem is strategy-proof when $|V|$ is odd.

Corollary 4.5. *The λ-1-median problem is strategy-proof when $|V|$ is odd.*

When combined, Theorem 4.2 and Corollaries 4.4 and 4.5 suggest that the only way to prevent manipulation is either to introduce bias or to hope that the number of agents is odd. However, we show that through careful mechanism design – by limiting agents' ideal points to a hyperrectangle – we can completely remove *the impact* of manipulation without introducing bias and without preventing manipulation.

Theorem 4.6. *Suppose $\mathcal{X} = \{x : a \leq x \leq b\}$ is a hyperrectangle. Then the price of anarchy of the λ-1-Median problem is 1, i.e., manipulation has no impact on the social cost of the outcome.*

Proof. To establish a price of anarchy of 1, it suffices to show that $r(\bar{\Pi}) \in [a(\Pi), b(\Pi)]$, i.e., the submitted outcome $r(\bar{\Pi})$ is a median for the sincere ideal points. As in Theorem 4.3, it suffices to show the result when $\mathcal{X}$ is the one-dimensional line segment $[a, b]$.

For contradiction and without loss of generality, assume $r(\bar{\Pi}) > b(\Pi)$. Since the set of sincere medians is given by the line segment $[a(\Pi), b(\Pi)]$, fewer than half of the agents have a sincere ideal point strictly more than $b(\Pi)$. However, since $r(\bar{\Pi})$ is a median for the submitted preferences, at least half of the agents submit an ideal point in the form $\bar{\pi}_v \geq r(\bar{\Pi})$. Therefore, there exists at least one agent v where $\pi_v \leq b(\Pi) < r(\bar{\Pi}) \leq \bar{\pi}_v$. Let $\bar{\pi}'_v = \bar{\pi}_v - \epsilon$ for some $\epsilon > 0$. Since $\pi_v < \bar{\pi}_v$, $\bar{\pi}'_v$ is more honest than $\bar{\pi}_v$ for sufficiently small ϵ. Further, if agent v moves their submitted ideal point to the left, then the set of medians either shifts to the left or does not move implying $r([\bar{\Pi}_{-v}, \bar{\pi}'_v]) \leq r(\bar{\Pi})$. Further, since $r(\cdot)$ is continuous, ϵ can be selected sufficiently small so that $\pi_v \leq r([\bar{\Pi}_{-v}, \bar{\pi}'_v]) \leq r(\bar{\Pi})$. This implies that v can obtain at least as good an outcome by submitting the more honest $\bar{\pi}'_v$. This contradicts minimal dishonestly implying that $r(\bar{\Pi}) \leq b(\Pi)$. A symmetric argument shows that $r(\bar{\Pi}) \geq a(\Pi)$. □

5 1-MEAN PROBLEM

In the 1-Mean problem, the outcome is selected by minimizing the social cost function $C(\Pi, x) := \sum_{v \in V} ||\pi_v - x||_2^2$. If the set of ideal points comes from a convex, compact set, then it is straightforward to show there is a unique minimizer and that the selection rule selects the outcome $r(\Pi) = \sum_{v \in V} \frac{\pi_v}{|V|}$. Since there is always a unique best response, every Nash equilibrium is also a minimally dishonest Nash equilibrium and we use the two terms interchangeably in this section.

For the 1-Mean problem, the impact of manipulation, i.e. the price of anarchy, depends on the underlying geometry of the domain $\mathcal{X}$. For arbitrary $\mathcal{X}$ the price of anarchy can be infinity, implying the 1-Mean selection rule has no power to guarantee outcomes close to the actual mean.

Theorem 5.1. *Suppose $p_v = 2$ for all $v \in V$. The price of anarchy of the 1-Mean problem in $\mathbb{R}^k$ for $k = 2$ is ∞ and when $|V| = 3$. Specifically, for all $\alpha \in (0, \pi/2)$ there exists an instance with a price of anarchy of $\frac{1}{2 \cdot \cos(\alpha)} \to \infty$ as $\alpha \to 0$.*

Proof. For $\alpha < \frac{\pi}{2}$, let $\bar{\pi}_1 = (0,0), \bar{\pi}_2 = (\frac{3}{2 \cdot \sin(\alpha)}, \frac{3}{2 \cdot \cos(\alpha)})$, and $\bar{\pi}_3 = (-\frac{3}{2 \cdot \sin(\alpha)}, \frac{3}{2 \cdot \cos(\alpha)})$, and let $\mathcal{X} = conv.hull(\bar{\pi}_1, \bar{\pi}_2, \bar{\pi}_3)$ as shown in Figure 3. Suppose the sincere ideal points are given by $\pi_1 = (0,0), \pi_2 = (\cos(\alpha), \sin(\alpha))$, and $\pi_3 = (-\cos(\alpha), \sin(\alpha))$. With respect to the sincere preferences, $C(\Pi, \vec{0}) = \sum_{v=1}^{3} ||\pi_v||_2^2 = 2$ implying the sincere cost of the optimal $r(\Pi)$ is at most 2.

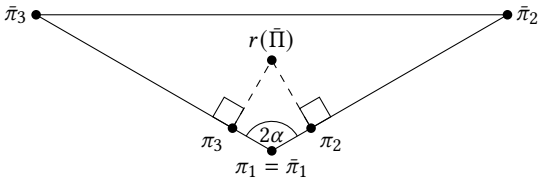

Figure 3: Preferences for Theorem 5.1. The sincere outcome will always have a total cost of 2 while the submitted outcome $r(\bar{\Pi})$ will move arbitrarily far away from $\vec{0}$ as $\alpha \to 0$ yielding a price anarchy of ∞.

Now consider the submitted 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 social cost of $C(\Pi, r(\bar{\Pi})) = \sum_{v=1}^{3} ||r(\bar{\Pi}) - \pi_v||_2^2 \geq ||r(\bar{\Pi}) - \pi_1||_2^2 = \frac{1}{\cos(\alpha)}$. If $\bar{\Pi}$ corresponds to a (minimally dishonest) Nash equilibrium of Π, then the price of anarchy is at least

$$\frac{C(\Pi, r(\bar{\Pi}))}{C(\Pi, r(\Pi))} \geq \frac{1}{2 \cdot \cos(\alpha)} \to \infty \text{ as } \alpha \to 0.$$

It remains to show that $\bar{\Pi}_v$ is a minimally dishonest best response for each agent. We first consider agent $v = 1$. Suppose agent 1 changes their preferences to $\bar{\pi}_1' \in \mathcal{X}$. After updating their preferences the outcome moves to $r([\bar{\Pi}_{-1}, \bar{\pi}_1']) = r(\bar{\Pi}) + \frac{1}{3}\bar{\pi}_1'$. This causes the outcome to move up and possibly to the left or the right. Regardless of p_1, this outcome is worse for agent 1 and therefore they are reporting a minimally dishonest best response.

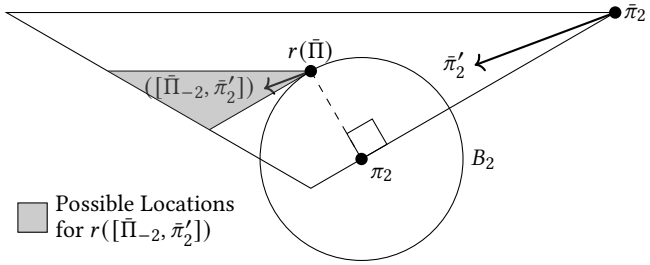

Figure 4: Possible locations for $r([\bar{\Pi}_{-2}, \bar{\pi}_2'])$.

Next, consider agent 2. Suppose agent 2 updates their preferences to $\bar{\pi}_2' = \bar{\pi}_2 - d$. Since $\bar{\pi}_2' \in \mathcal{X}$, $d \in cone((1,0), (\pi_2))$. After updating their preferences, the selected outcome will be $r([\bar{\Pi}_{-2}, \bar{\pi}_2']) = r(\bar{\Pi}) - \frac{1}{3} \cdot d$. Let $B_2 = \{x : ||x - \pi_2||_2 \leq ||\pi_2 - r(\bar{\Pi})||_2$ be the set of ideal points that agent 2 prefers to $r(\bar{\Pi})$. By construction, $\{x \in \mathbb{R}^2 : \pi_2^\intercal x = \pi_2^\intercal r(\bar{\Pi})\}$ is tangent to B_2 as shown in Figure 4. Thus, for all $d \in cone((1,0), (\pi_2))$ where $||d|| > 0$, reporting $\bar{\pi}_2 - d$ would yield a worse outcome for agent 2 and agent 2 is minimally dishonest. A symmetric argument implies agent 3 is minimally dishonest and $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. Thus, the price of anarchy is at least $\frac{1}{2 \cdot \cos(\alpha)} \to \infty$ as $\alpha \to 0$. □

We remark that Theorem 5.1 can be generalized for arbitrary $p_v \geq 1$ by carefully constructing the triangle so that $r(\bar{\Pi})$ is tangent to the l_{p_v}-ball at π_2. Further, the result can be generalized for an arbitrary number of agents by increasing the height of the triangle and by adding agents with ideal points at $\vec{0}$.

Like the 1-Median problem, through better mechanism design, the impact of manipulation can be reduced. The 1-Mean selection rule is better behaved when the set of ideal points is limited to a hyperrectangle. Specifically, the submitted outcome will always be in the smallest bounding box containing all sincere ideal point, resulting in a price of anarchy that is $O(|V|)$. Thus, the impact of manipulation is limited to being linear in the number of agents.

Lemma 5.2. *Suppose $\mathcal{X} = \{x \in \mathbb{R}^k : a \leq x \leq b\}$. Let $l(\Pi) \in \mathbb{R}^k$ and $u(\Pi) \in \mathbb{R}^k$ be such that $\{x \in \mathbb{R}^k : l(\Pi) \leq x \leq u(\Pi)\}$ is a box containing Π. Then for every (minimally dishonest) Nash equilibrium $\bar{\Pi}$, $l(\Pi) \leq r(\bar{\Pi}) \leq u(\Pi)$.*

Proof. By symmetry, it suffices to show that $r(\bar{\Pi}) \geq l(\Pi)$. For contradiction, suppose $\bar{\Pi}$ is a Nash equilibrium such that $r_i(\bar{\Pi}) < l_i(\Pi)$ for some $i \in [k]$. This implies there exists an agent v such that $\bar{\pi}_{vi} \leq r_i(\bar{\Pi}) < l_i(\Pi)$ where $\pi_v = (\pi_{v1}, \pi_{v2}, ..., \pi_{vk})$.

Since $\bar{\pi}_{vi} < l_i(\Pi)$ and since $\mathcal{X}$ is a hyperrectangle, $\bar{\pi}_v' = \bar{\pi}_v + \epsilon \cdot e_i \in \mathcal{X}$ for sufficiently small ϵ where e_i is the ith standard basis vector. This shift in $\bar{\pi}_v$ causes the outcome to shift to $r([\bar{\Pi}_{-v}, \bar{\pi}_v']) = r(\bar{\Pi}) + \frac{\epsilon}{|V|} \cdot e_i$. Further, since $r_i(\bar{\Pi}) < l_i(\bar{\Pi})$, ϵ can be selected small enough so that $r_i(\bar{\Pi}) < r_i([\bar{\Pi}_{-v}, \bar{\pi}_v']) < l_i(\Pi) \leq \pi_{vi}$. Agent v's associated cost for this new outcome is

$$\left(\sum_{j \in [k]} (r_j([\bar{\Pi}_{-v}, \bar{\pi}_v']) - \pi_{vj})^{p_v} \right)^{1/p_v}$$

$$= \left((r_i([\bar{\Pi}_{-v}, \bar{\pi}'_v]) - \pi_{vi})^{p_v} + \sum_{j \neq i} (r_j(\bar{\Pi}) - \pi_{vj})^{p_v} \right)^{1/p_v}$$
$$< \left((r_i(\bar{\Pi}) - \pi_{vi})^{p_v} + \sum_{j \neq i} (r_j(\bar{\Pi}) - \pi_{vj})^{p_v} \right)^{1/p_v}$$
$$= \left(\sum_{j \in [k]} (r_j(\bar{\Pi}) - \pi_{vj})^{p_v} \right)^{1/p_v}$$

since $r_i(\bar{\Pi}) < r_i([\bar{\Pi}_{-v}, \bar{\pi}'_v]) < \pi_{vi}$. Therefore agent v prefers the new outcome obtained by submitting $\bar{\pi}'_v$ contradicting that $\bar{\Pi}$ is a Nash equilibrium. Thus $l(\pi) \leq r(\bar{\Pi}) \leq u(\pi)$. □

Theorem 5.3. *Suppose $\mathcal{X} = \{x \in \mathbb{R}^k : a \leq x \leq b\}$. Then the price of anarchy of the 1-Mean problem in $\mathbb{R}$ is between $|V|$ and $2 \cdot |V|$.*

Proof. First we show an upper bound of $2 \cdot |V|$. Let $\bar{\Pi}$ be a Nash equilibrium for the sincere profile Π. Let $[l, u]$ be the smallest box such that $\Pi \in [l, u]$. Without loss of generality, we may assume $l = \vec{0}$. This implies that for each axis $i \in [k]$ that there exists agents v and v' such that $\pi_{vi} = u_i$ and $\pi_{v'i} = 0$

The sincere outcome is $r(\Pi) = \sum_{v \in V} \frac{\pi_v}{|V|}$ and $r(\Pi) \in [0, u]$. The cost of the sincere outcome $r(\Pi)$ is then

$$\sum_{v \in V} ||\pi_v - r(\Pi)||_2^2$$
$$= \sum_{i \in [k]} \sum_{v \in V} (\pi_{vi} - r_i(\Pi))^2$$
$$\geq \sum_{i \in [k]} \left((\min\{\pi_{vi}\} - r_i(\Pi))^2 + (\max\{\pi_{vi}\} - r_i(\Pi))^2 \right)$$
$$= \sum_{i \in [k]} \left(r_i(\Pi)^2 + (u_i - r_i(\Pi))^2 \right)$$
$$\geq \sum_{i \in [k]} \left(\left(\frac{u_i}{2}\right)^2 + \left(u_i - \frac{u_i}{2}\right)^2 \right) = \sum_{i \in [k]} \frac{u_i^2}{2} = ||u||_2^2/2$$

since $f(x) = x^2 + (w - x)^2$ is minimized at $x = w/2$.

Now consider an equilibrium $\bar{\Pi}$. By Lemma 5.2, $r(\bar{\Pi}) \in [0, u]$ and every sincere ideal point π_v is at most $||u||_2^2$ away from $r(\bar{\Pi})$. Thus, the sincere cost of $r(\bar{\Pi})$ is at most $|V| \cdot ||u||_2^2$ and the manipulation causes the social cost to increase by a factor of at most

$$\frac{\sum_{v \in V} ||\pi_i - r(\bar{\Pi})||_2^2}{\sum_{v \in V} ||\pi_i - r(\Pi)||_2^2} \leq \frac{|V| \cdot ||u||_2^2}{\frac{||u||_2^2}{2}} = 2 \cdot |V|.$$

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

Now consider the submitted preferences $\bar{\Pi}$ where $\bar{\pi}_1 = \vec{1} \cdot |V|$ and $\bar{\pi}_v = \pi_v$ for all other v. With respect to these preferences, the selected outcome is $r(\bar{\Pi}) = \vec{1}$ for a sincere social cost of $\sqrt{k} \cdot (|V|-1)$. 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 anarchy is between $|V|$ and $2 \cdot |V|$ completing the proof of the theorem. □

6 MINIMIZING l_2 NORM

Finally, we consider the social cost function $C(\Pi, x) = \sum_{v \in V} ||\pi_v - x||_2$. In 1-dimension, minimizing the l_1 norm (1-median) is equivalent to minimizing the l_2 norm, and our results from Section 4 extend to the l_2 norm when $\mathcal{X} \subset \mathbb{R}$. However, in higher dimensions, the problems are drastically different. We show that if ties are broken when selecting the midpoint, then the price of anarchy is infinity even if $\mathcal{X}$ is a rectangle.

Theorem 6.1. *Let $\mathcal{X} = \{x \in \mathbb{R}^2 : a \leq x \leq b\}$ and suppose $|V| \geq 4$ is even. When individuals are minimally dishonest and ties are broken by selecting the center point, the price of anarchy when minimizing the sum of l_2 norm is ∞. Specifically, for all $\epsilon \in (0, 1]$ there exists an instance with a price of anarchy of $\frac{|V|-\epsilon}{\epsilon} \to \infty$ as $\epsilon \to 0$.*

Proof. We may assume $\{(0, 0), (-1, 1), (1, 1)\} \in \mathcal{X}$ by translating $\mathcal{X}$. Let $|V| \geq 4$ be even and suppose $\pi_v = (0, \epsilon)$ for an arbitrary voter and that $\pi_v = (0, \epsilon)$ for all other v. If everyone is honest then the outcome is $r(\Pi) = (0, 0)$ with a cost of ϵ. Consider $\bar{\Pi}$ where half of the agents submit $(-1, 1)$ and the other half submit $(1, 1)$. With respect to $\bar{\Pi}$, $(x, 1)$ minimizes $C(\bar{\Pi}, x)$ for all $x \in [-1, 1]$ resulting in $r(\bar{\Pi}) = (0, 1)$. However, with respect to Π, the social cost is $|V| - \epsilon$.

We now show $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. By symmetry, it suffices to examine agent v that submits $\bar{\pi}_v = (-1, 1)$. If agent 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 agent v. If $\bar{\pi}'_v = (w, 1)$ for $w < -1$, then the outcome does not change and agent v is less honest. If agent v submits $\bar{\pi}'_v = (w, 1)$ for some $w \in (-1, 1)$, then the new outcome is $(\frac{1+w}{2} \geq 0, 1)$, which is worse for agent v. Finally, if agent v submits $\bar{\pi}'_v = (w, 1)$ for $w > 1$, then $r([\bar{\Pi}_{-1}, \bar{\pi}'_1]) = (1, 1)$ yielding a worse outcome for agent v. All possibilities yield either worse outcomes for agent v or cause agent v to be less honest without any benefit. Thus agent v is giving the unique minimally dishonest best response and $\bar{\Pi}$ is a minimally dishonest Nash equilibrium. Therefore, the price of anarchy when minimizing the l_2 norm is at least $\frac{|V|-\epsilon}{\epsilon} \to \infty$. □

7 CONCLUSION

We have shown that the price of anarchy is a powerful tool for measuring the impact of manipulation. Further, we have shown that not all forms of manipulation are equal; through careful mechanism design, we have uncovered a class of spatial social choice selection rules that are immune to negative consequences of manipulation despite remaining manipulable. This is in contrast to standard approaches that sacrifice other beneficial properties, e.g., unbiased tie-breaking, to gain strategy-proofness. This new approach opens up many new possibilities for the design of social choice mechanisms in general and we propose that the price of anarchy be one of the criteria by which a selection rule is assessed.

ACKNOWLEDGMENTS

Our research has been supported by NSF under grant number CMMI-1335301. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the sponsoring organizations, agencies, or governments.

REFERENCES

[1] Nir Ailon, Moses Charikar, and Alantha Newman. 2008. Aggregating Inconsistent Information: Ranking and Clustering. *jacm* 55, 5 (2008), 23:1–23:27. https://doi.org/10.1145/1411509.1411513

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

[3] James P Bailey and Craig A Tovey. [n.d.]. The Price of Deception in Spatial Social Choice. ([n. d.]).

[4] James P Bailey and Craig A Tovey. 2021. Conditions for Stability in Strategic Matching. *arXiv preprint arXiv:2108.04381* (2021).

[5] James P Bailey and Craig A Tovey. 2024. Impact of Tie-Breaking on the Manipulability of Elections. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. 105–113.

[6] John Bartholdi, Thomas Seeley, Craig Tovey, and John VandeVate. 1993. The Pattern and Effectiveness of Forager Allocation among Food Sources in Honey Bee Colonies. *jtb* 160 (1993), 23–40.

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

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

[9] O. Davis, M. Degroot, and M. Hinich. 1972. Social Preference Orderings and Majority Rule. *Econometrica* 40 (1972), 147–157.

[10] Elad Dokow and Dvir Falik. 2012. Models of Manipulation on Aggregation of Binary Evaluations. *CoRR* abs/1201.6388 (2012). http://arxiv.org/abs/1201.6388 Presented at COMSOC 2012.

[11] Elad Dokow and Ron Holzman. 2010. Aggregation of binary evaluations. *Journal of Economic Theory* 145, 2 (2010), 495 – 511. https://doi.org/10.1016/j.jet.2007.10.004 Judgment Aggregation.

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

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

[14] Bhaskar Dutta and Jean-François Laslier. 2010. Costless Honesty in Voting. In *Proceedings of Social Choice and Welfare.*

[15] Bhaskar Dutta and Arunava Sen. 2012. Nash implementation with partially honest individuals. *Games and Economic Behavior* 74, 1 (2012), 154 – 169. https://doi.org/10.1016/j.geb.2011.07.006

[16] Bruno Escoffier, Laurent Gourvès, Nguyen Kim Thang, Fanny Pascual, and Olivier Spanjaard. 2011. Strategy-proof Mechanisms for Facility Location Games with Many Facilities. In *Proceedings of the Second International Conference on Algorithmic Decision Theory* (Piscataway, NJ) *(ADT'11)*. Springer-Verlag, Berlin, Heidelberg, 67–81. http://dl.acm.org/citation.cfm?id=2050843.2050849

[17] Dimitris Fotakis and Christos Tzamos. 2013. Strategy-Proof Facility Location for Concave Cost Functions. *CoRR* abs/1305.3333 (2013). http://arxiv.org/abs/1305.3333

[18] Uri Gneezy. 2005. Deception: The Role of Consequences. *The American Economic Review* 95, 1 (2005), 384–394. http://www.jstor.org/stable/4132685

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

[20] Sjaak Hurkens and Navin Kartik. 2009. Would I lie to you? On social preferences and lying aversion. *Experimental Economics* 12, 2 (2009), 180–192. http://EconPapers.repec.org/RePEc:kap:expeco:v:12:y:2009:i:2:p:180-192

[21] Navin Kartik, Olivier Tercieux, and Richard Holden. 2014. Simple mechanisms and preferences for honesty. *Games and Economic Behavior* 83, C (2014), 284–290. https://EconPapers.repec.org/RePEc:eee:gamebe:v:83:y:2014:i:c:p:284-290

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

[23] G. Kramer. 1972. Sophisticated Voting over Multidimensional Spaces. *Journal of Mathematical Sociology* 2 (1972), 165–180.

[24] Jean-François Laslier, Matías Núñez, and Carlos Pimienta. 2017. Reaching consensus through approval bargaining. *Games and Economic Behavior* 104 (2017), 241 – 251. https://doi.org/10.1016/j.geb.2017.04.002

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

[26] Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Allen Zhu. 2010. Asymptotically Optimal Strategy-proof Mechanisms for Two-facility Games. In *Proceedings of the 11th ACM Conference on Electronic Commerce* (Cambridge, Massachusetts, USA) *(EC '10)*. ACM, New York, NY, USA, 315–324. https://doi.org/10.1145/1807342.1807393

[27] Hitoshi Matsushima. 2008. Role of honesty in full implementation. *Journal of Economic Theory* 139, 1 (2008), 353 – 359. https://doi.org/10.1016/j.jet.2007.06.006

[28] R. McKelvey. 1976. Intransitivities in Multi-dimensional Voting Models and some Implications for Agenda Control. *Journal of Economic Theory* 12 (1976), 472–482.

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

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

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

[32] Matías Núñez and Jean-François Laslier. 2015. Bargaining through Approval. *Journal of Mathematical Economics* 60 (2015), 63 – 73. https://doi.org/10.1016/j.jmateco.2015.06.015

[33] Svetlana Obraztsova, Omer Lev, Evangelos Markakis, Zinovi Rabinovich, and Jeffrey S. Rosenschein. 2017. Distant Truth: Bias Under Vote Distortion Costs. In *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems* (São Paulo, Brazil) *(AAMAS '17)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 885–892. http://dl.acm.org/citation.cfm?id=3091125.3091250

[34] Svetlana Obraztsova, Evangelos Markakis, and David R. M. Thompson. 2013. Plurality Voting with Truth-Biased Agents. In *Algorithmic Game Theory*, Berthold Vöcking (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 26–37.