Voting systems for multiple choice Schelling games (Draft)

Abstract

In this work we consider social choice problems motivated by blockchain oracles. Here participants are asked to respond to semi-objective questions and are economically incentivized to provide responses that agree with each other, with the idea that the true response is then a "Schelling point". We define a notion of honest behaviour in this framework and then show an impossibility theorem that demonstrates that, under minimal assumptions on the underlying voting rule and payoff structure chosen, if there are three or more possible options being voted upon, there will be situations where participants are incentivized to participate "dishonestly".

1 Introduction

In recent years blockchains have offered new possibilities for economic relationships in the absence of trust between parties. However, useful applications often need access to information that is not verified as part of the internal blockchain consensus protocol. For example, a blockchain-based prediction market platform might need to know which candidate won an election in order to make payouts afterwards. Similarly, a crop insurance contract might need to know whether it rained in Saskatchewan. The problem of making such information available to blockchain applications without adding additional trust assumptions on parties, is known as the "blockchain oracle problem".

One solution to this problem has been to make use of Schelling games [25], [28]. The idea of this structure is to set up a coordination game [9], namely a game in which participants are asked to vote on the answer to some external question and they are rewarded or penalized based on whether they agree with the votes of others. Then, in many contexts, one expects that the true answer will be a Schelling point (or focal point) in the sense of Thomas Schelling [22], namely "a solution that people will tend to use [in attempts to coordinate] in the absence of communication, because it seems natural, special, or relevant to them".

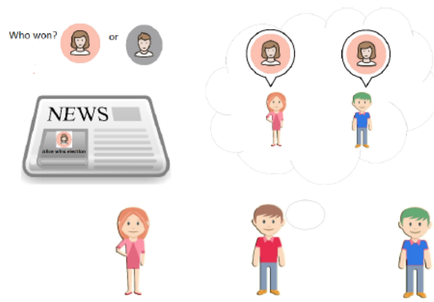

Figure 1: Each voter expect the other voters to submit the "honest" outcome as it is a distinguished choice. Hence as the voters are incentivized to be coherent, they have an incentive to vote honestly.

	X wins	Y wins
$\mathcal{USR}$ votes X	p	0
$\mathcal{USR}$ votes Y	0	p

Figure 2: An example payoff matrix for a voter $\mathcal{USR}$ in a symmetrical, binary coordination game. If one of X or Y is "true", then one might expect that response to be a Schelling point for voters.

One might attempt to generalize this approach to semi-objective questions with more than two possible outcomes. For example, a decentralized, blockchain-based marketplace might use a Schelling game system to determine whether a seller respected the conditions of a sale with possible outcomes of no refund, a partial refund, or a full refund to the buyer. A decentralized social media platform might use a Schelling game to determine whether a post violated its terms and conditions, with possbile outcomes including the post remaining available, the post being removed but the user who posted it not being banned, or the post being removed and the user who posted it being banned [15].

This framework raises a number of interesting research questions. Particularly, one asks what voting rules should be used to convert user responses into a collective outcome, and what payoff structure should be put in place to determine the rewards and penalties for different responses. For semi-objective questions, one might take the point of view that there exists some "true" response to the question that voters see with noise. This is similar to the approach of the analysis of [8]. Already in that work interesting questions are raised with respect to what types of voting rules can have desirable properties in aggregating user responses so that the collective outcome is likely to be the underlying "true" value; this has natural applications to the design of voting rules for crowdsoucring or micro-task platforms. However, in [8], the incentivization of voters is not considered, rather they are assumed to report the noisy truth as they observe it. Such a perspective is not sufficient for the adversarial setting of blockchain applications. Hence, in our setting, we are interested in the degree to which it is possible to choose a voting rule and payoff structure that both *actually* incentivize participants to submit votes that reflect the noisy truth as they observe it *and* then aggregate these votes effectively. Such attempts to incentivize particualr voting behaviour in participants raises parallels to some of the impossibility theorems that have been seen in standard elections [3], [12], [21].

Note that despite the presence of third parties who can (potentially) moderate non-blockchain crowdsourcing platforms, these services nevertheless have issues regarding spam responses [2], [14]. Thus, the framework of Schelling game based mechanisms discussed in this work could potentially be implemented for semi-objective questions on such platforms in an effort to improve response quality.

2 Related work

This work, as it ultimately deals with mechanisms by which the information provided by many individuals can be collected into a collective choice, draws on a wide variety of ideas in social choice theory. In traditional elections, one typically imagines voters as having preferred choices; namely that they receive more or less utility depending on which candidate is chosen by the group. Then the motivation of participants is to vote in such a way that the resulting collective choice will provide them with as much utility as possible. Already this problem gives rise to a rich set of complexities; particularly, it has been seen that many natural properties that one would want electoral systems to have are incompatible [3], [12], [21], [17]. Remark, for our purposes that the classic impossibility theorems of social choice theory such as Arrow's Impossibility Theorem [3] and the Gibbard-Satterthwaite Theorem [12], [21], to the degree that they make statements about the positions of options in submitted lists, still apply directly in a setting where the voters have another incentive structure, while they no longer necessarily have the same implications in terms of incentive compatibility. Also, of particular note for our purposes is work in which voters are rewarded or penalized financially based on how they vote, which is the perspective of [24] which reconsiders the Gibbard-Satterthwaite Theorem in the context where one allows for divisible private goods (money) to be distributed among the voters in a way that is determined by the outcome of the vote.

An alternative framework, relevant to the design of crowdsourcing platforms, is to suppose that there exists some "truth" that voters observe with noise. This perspective already dates back to Condorcet's Jury Theorm [10]. Then one can pose questions about which voting rules, if any, have good properties in terms of converting the information provided by voters into collective outcomes that are likely to recover the "truth". Theoretical work along these lines is considered in [8], [19], [7], and behavioural experiments are considered in

[16]. However, this perspective generally assumes that voters honestly report the truth as they observe it. This point of view is not realistic in the adversarially, pseudo-anonymous environment of blockchain applications. Existing work that reintroduces strategic voting and individual incentives into the perspective of this model is [4] where it is noted that Condorcet's Jury Theorem implicitly assumes that votes reflect the truth as observed by participants, and indeed the strategy of providing such a vote does not generally give a Nash equilibrium. In [13], results along the lines of the Jury Theorem are recovered with potentially strategic voters in the context of a set of options that is strictly ordered. However, in these works voters' incentives when considering to cast a strategic vote come from an instrinsic desire to have the "true" outcome (as well as they are able to evaluate it through the noise) adopted.

Meanwhile, many prominent blockchain oracles [25], [28], [1], [18], [15] rely on some version of a Schelling game idea. In this setting, participants are generally incentivized purely through the payouts of the game and only have a weak intrinsic interest in the system producing honest results. Note that in a more general sense, Schelling (or focal) points have a long tradition in social choice theory as a model for collaboration [23].

Past work considering such Schelling games often focuses on situations where voters are presented with a binary choice (see for example [1]), avoiding both the classical problems of social choice theory in multi-candidate elections. To the degree that systems such as [25] do present voters with non-binary choices, the "honest" choice is typically so clear (i.e. reporting the winner of a well-known election), that it is not problematic to choose the winner via a Plurality voting system. In contrast, our current work is particularly motivated for applications to Kleros [15], a blockchain-based dispute resolution platform. Kleros, as a sort of "dispute resolution" oracle, operates under similar principles to other Schelling game systems, but it is designed to be able to handle cases where there is less likely to be unanimity around winning choices; in these cases, use of a Plurality voting system may be less appropriate in the presence of non-binary outcomes.

Finally in [11], a Schelling point based system that allows for non-binary answers, specifically that allows for real value answers, is considered, but this is done by binarizing the choices provided to voters. (Note that taking a binarization approach to multiple outcome decisions more generally has limitations, compare to the discussion on discursive dilemmas in 17.1 of [5].)

3 Contribution of this work

In this work we will explore questions of incentive compatibility in multiple choice Schelling games. Particularly, in Section 5 we define a natural concept of "Schelling-honesty" for what an "honest vote" might mean in this context. In Section 7, under very minimal assumptions on voting and payoff systems, we prove in Theorem 1 that there will always be some situations where participants will be incentivized to deviate from Schelling-honesty. This theorem draws heavily upon the Gibbard-Satterthwaite Theorem, and could naturally be seen as an analog of that result in this context. In Section 8 we show that these (minimal) assumptions on the voting system and payoff system in Theorem 1 are necessary by providing counterexamples when these criteria are not satisfied. In Section 9, we consider some specific examples where Schelling-honest voting is not economically optimal, and we explore the question of under what conditions on a voting system is it always possible to even provide a Schelling-honest vote. In Section 10, we consider how this work might fit with broader goals in designing Schelling games with multiple choices.

4 Notation and basic notions

We use the following notation throughout: A is a set of possible outcomes for a voting system. $\mathcal{L}$ is the set of possible orderings of the outcomes in A. N is the number of voters. We denote the number of options by $\#A = n$. A social choice function is a function $f : \mathcal{L}^N \to A$. Often we will consider the perspective of a single voter $\mathcal{USR}_i$. Then if we denote $\mathcal{USR}_i$'s vote by $v_i \in \mathcal{L}$, we denote the votes of the other users by $v_{-i} \in \mathcal{L}^{N-1}$, and one can compute the outcome given these votes as $f(v_i, v_{-i})$.

Note that for a permutation σ on the set of outcomes A, one has natural maps on $\mathcal{L}$, $\mathcal{L}^{N-1}$, and $\mathcal{L}^N$ given by applying σ to each element of each ranked list. We denote these maps also by σ when doing so does not create ambiguity.

Then, it is natural to adapt certain ideas from standard social choice theory. Of particular relevance are the following concepts:

Definition 1. *A social choice function f is said to be neutral if for any permutation σ on A and any $l_1 \dots, l_N \in \mathcal{L}$,*

$$f(\sigma(l_1), \dots, \sigma(l_N)) = f(l_1, \dots, l_N).$$

Definition 2. *A social choice function f is said to have the absence of veto (AV) property if, for any choice $a \in A$, if a is at the top of all but at most one rankings l_i, then $f(l_1, \dots l_N) = a$.*

Definition 3. *A social choice function f is said to be monotonic if for any $l_1 \dots, l_N, l'_j \in \mathcal{L}$ where $f(l_1, \dots, l_j, \dots, l_N) = w$ and w is ranked higher according to l'_j than it is according to l_j, then $f(l_1, \dots, l'_j, \dots, l_N) >_{l_j}= w$.*

We will consider as needed several other standard concepts of social choice functions such as dictators, the Condorcet criterion, Condorcet winners, and Pareto efficiency. See, for example, [5] for definitions of these properties. Furthermore, while it is possible that there are new, not yet considered, voting systems that are particularly adapted to the context of Schelling games, most of the examples that we consider in this work are drawn from standard voting systems (combined with some payoff system) such as Plurality, Instant-Runoff, and Ranked Pairs; see [5] for descriptions of these systems. In this work, our results will apply to resolute voting systems, namely on any given input of $\mathcal{L}^N$

the corresponding social choice function returns some outcome in A even in the event of a tie. Hence, when we consider the above mentioned standard voting systems, we will assume that there is some unspecified determininstic tie-break system being used.

Finally, we will have a notion of a payoff system; namely to each voter $\mathcal{USR}_i$ will be paid an amount given by a function

$$G_i : \mathcal{L}^N \to \mathbb{R}.$$

Considerations around payoff systems will be further discussed in Section 6.

5 Model

We consider the following model. A given voter $\mathcal{USR}_i$ has some probabilistic expectations for the behaviour of the other voters. We formalize this by attaching to each voter a probability function that we call her *evaluation*:

$$V_{-i} : \mathcal{L}^{N-1} \to [0, 1].$$

The other voters' rankings are not necessarily identically distributed. For example, $\mathcal{USR}_i$ might know exactly how a given other voter will vote with probability one while only having a probabilistic knowledge of others. Nor are the other voters' rankings necessarily independent of each other. Indeed $\mathcal{USR}_i$'s evaluation for their votes could include knowledge that voter j_1 will always vote exactly the same way as voter j_2 even if $\mathcal{USR}_i$ does not know either of their votes in advance. However, we assume that all of the other voters' rankings are independent of $\mathcal{USR}_i$'s vote. Then, for any given vote by $\mathcal{USR}_i$, $\mathcal{USR}_i$ can consider the probability that V_{-i} take each given value and she can thus estimate the probability of each outcome.

We consider participants that are economically rational in the sense that they take actions that maximize their expected return, where the uncertainty in their return is captured in their evaluations. In the context of questions where there is some "true" answer that participants see with noise, participants ideally should be incentivized to cast votes that reflect their honest assessement. Then, the basic idea of a Schelling game is that participants' expectations of the winning outcome of the vote should be used by the incentive system as a proxy for this "true" answer. Indeed, if voters believe that other participants will arrive at similar assessements as themselves, they might believe that there will be an equilibrium where providing a ranked ordering of the options in order of the probability that they have of being the "true" answer (hence providing votes such as those considered in [8]) is a good proxy for a ranking of the options in order of their probability of winning. How good this proxy is, namely how closely the ranking of options in order of winning aligns with the ranking of options in order of being "true", is then a question that builds on the viewpoint of voting systems as stastical estimators [8], [7], [19].

Then, one asks to what degree it is, in fact, possible to incentivize voters to pursue this “Schelling” strategy of ranking options by their chances of winning. This motivates us to define “honest” behaviour in our context as the following:

Definition 4. *Given an estimation V_{-i} for how a user believes that other voters will vote which is independent from her own vote, $\mathcal{USR}_i$'s vote $v_i = a_1 > a_2 > \ldots > a_M$ is said to be Schelling-honest if*

$$p(a_1) \geq p(a_2) \geq \ldots \geq p(a_M),$$

where $p(a_j) = prob(a_j \text{ wins} | \mathcal{USR}_i \text{ votes } v_i)$ is the probability of a_j winning based on V_{-i} and assuming that $\mathcal{USR}_i$ votes $v_i = a_1 > a_2 > \ldots > a_M$.

Note that there may be several Schelling-honest votes possible to a user, to the degree that the user's vote changes the ordering of the most likely outcomes to win. Regardless, it is very natural to consider a vote which is not in order by probability as *dishonest*. We consider the degree to which a Schelling game incentive system encourages voters to submit such votes as a failure of incentive compatibility.

Similarly, we define

Definition 5. *An evaluation V_{-i} is said to be Schelling-dishonest if the vote that yields the highest payoff for $\mathcal{USR}_i$ in according to this evaluation is not Schelling-honest.*

6 Payoff systems

We begin this section with a natural criterion that a payoff system might satisfy. Then we discuss several examples of payoff systems from the perspective of this criterion.

Definition 6. *A payoff system that pays to $\mathcal{USR}_i$ an amount given by*

$$G_i : \mathcal{L}^N \to \mathbb{R}$$

is said to be reasonable if:

- *Suppose $v_{-i} \in \mathcal{L}^{N-1}$ is some fixed set of choices for all votes other than those of $\mathcal{USR}_i$. Suppose r_i, $s_i \in \mathcal{L}$ are such that the winner $f(r_i, V_{-i})$ is ranked in the k_1st place of r_i and the winner $f(s_i, V_{-i})$ is ranked in the k_2nd place of s_i. Then if, $k_1 \leq k_2 \Rightarrow G_i(r_i, V_{-i}) \geq G_i(s_i, V_{-i})$.*

- *For any $v_i \in \mathcal{L}$, $V_{-i} \in \mathcal{L}^{N-1}$ and permutation σ of the options, if $f(v_i, V_{-i}) = a$ and $f(\sigma(v_i), \sigma(V_{-i})) = \sigma(a)$, then $G(v_i, V_{-i}) = G(\sigma(v_i), \sigma(V_{-i}))$.*

Furthermore, G_i is said to be strictly reasonable if:

- *G_i is reasonable.*

- *Suppose r_i, $s_i \in \mathcal{L}$ are such that the winner $f(r_i, v_{-i})$ is ranked in the k_1st place of r_i and the winner $f(s_i, v_{-i})$ is ranked in the k_2nd place of s_i. Then if, $k_1 < k_2 \Rightarrow G_i(r_i, v_{-i}) > G_i(s_i, V_{-i})$.*

When the relevant voter $\mathcal{USR}_i$ is clear from context, we will sometimes simply denote $G_i = G$.

In all cases, we suppose that all voters submit an ordering in $\mathcal{L}$ along with a deposit D. (Hence the minimal possible payoff for a given vote is $-D$.)

Definition 7. *(All or nothing incentive system)*

- *Determine the winner $w \in A$ via an underlying voting system.*
- *Any voter that places w first in her list is returned her deposit D. All other deposits are burned.*

Definition 8. *(Penalty based incentive system)*

- *Determine the winner $w \in A$ via an underlying voting system.*
- *For each pair of the winner versus another choice in A, e.g. w versus a, w versus b, etc any voter who did not rank w ahead of the other choice loses a deposit $d = \frac{D}{\#A-1}$.*
- *These lost deposits are burnt.*

Definition 9 (Redistributive payoff system)**.**

- *Determine the winner $w \in A$ via an underlying voting system.*
- *For each pair of the winner versus another choice in A, e.g. w versus a, w versus b, etc any voter who did not rank w ahead of the other choice loses a deposit $d = \frac{D}{\#A-1}$.*
- *Voters are paid a reward for their coherent votes of*

$$\frac{\#\ \textit{total incoherent votes across all pairs}}{\#\ \textit{total coherent votes across all pairs}} \#\ \textit{pairs in which}\ \mathcal{USR}\ \textit{coherent} \cdot d.$$

Similarly, a platform can provide a total net reward R that is distributed to voters according to

$$\frac{R}{\#\ \textit{total coherent votes across all pairs}} \#\ \textit{pairs in which}\ \mathcal{USR}\ \textit{coherent.}$$

Proposition 1. *The all or nothing, penalty based, and redistributive payoff systems are all reasonable. Furthermore, the penalty based and redistributive payoff systems are strictly reasonable.*

Proof. For all three systems, suppose σ is a permutation of A, $v_i \in \mathcal{L}$, $v_i \in \mathcal{L}^{N-1}$, such that $f(v_i, v_{-i}) = w$ and $f(\sigma(v_i), \sigma(v_{-i})) = \sigma(w)$. Then, for the all or nothing system and the penalty based system, the position of $\sigma(w)$ in $\sigma(v_i)$ is the same as the position of w in v_i. So these systems give the same payout after permutation. Moreover, when computing $G(\sigma(v_i), \sigma(v_{-i}))$, as $\sigma(w)$ has the same place in each participants vote under $\sigma(v_{-i})$ as w does under v_{-i}, the number of coherent and incoherent pairs is the same before and after the permuation. So the redistributive payoff system also preserves payouts under σ.

For the all or nothing and the penalty based payoff systems, the payoff monotonically decreases with the position that $\mathcal{USR}_i$ places w by fixed amounts regardless of v_{-i}. For the penalty based payoff system this amount is non-zero in each place. For the redistributive system, note that for a fixed v_{-i}, if $\mathcal{USR}_i$ places w in the k_ith place, we can compute

$$\frac{\#\text{ total incoherent votes across all pairs}}{\#\text{ total coherent votes across all pairs}} \#\text{ pairs in which } \mathcal{USR} \text{ coherent} \cdot d$$

$$= \frac{\text{const}_1 + (k_i - 1)}{\text{cont}_2 + (\#A - k_i + 1)} \cdot (\#A - k_i + 1) \cdot d$$

which is strictly decreasing in k_i.

□

7 An impossibility result

We see that, under very limited assumptions on a social choice function and payoff structure, that inevitably there will be situations where participants are incentivized to deviate from the notion of Schelling-honesty that we have described in Section 5.

Theorem 1. *Consider $k = \#A \geq 3$. Suppose $f : \mathcal{L}^N \to A$ is a neutral social choice function satisfying absence of veto (AV). Then, for any strictly reasonable payoff structure, there exists a Schelling-dishonest evaluation.*

Note that the voting system is not assumed to be anonymous i.e. that it is symmetric under permutations of the voters[1].

In the remainder of this section we develop the tools allowing us to eventually prove this result.

Lemma 1. *Suppose f is reasonable. Take $v, v' \in \mathcal{L}$ and a probability function V_{-i} for the votes of other users. Then*

$$E[G(v, V_{-i}) - G(v', V_{-i})] = \sum_{\substack{V_{-i} = x \\ v, v' \text{ differ on} \\ \text{place of winner} \\ \text{for } x}} prob(V_{-i} = x) \left(G(v, V_{-i}) - G(v', V_{-i}) \right).$$

[1]Indeed, for our motivating application to Kleros [15], there are situations where one participant has many tokens and is drawn more than once for the same case, creating an asymmetry between the voters.

Proof. We compute

$$E[G(v, V_{-i}) - G(v', V_{-i})] = \sum_{V_{-i}=x} \text{prob}(V_{-i} = x) \left(G(v, V_{-i}) - G(v', V_{-i})\right)$$

Consider any $a \in A$ such that a is in the jth position for each of v, v' and take some fixed value x for V_{-i} such that $f(v, x) = f(v', x) = a$. Then, by applying the inequality of 6 in both directions, $G(v, x) = G(v', x)$. Hence all terms corresponding to $V_{-i} = x$ for which v and v' agree on the position of the winner cancel, and we are left with

$$E[G(v, V_{-i}) - G(v', V_{-i})] = \sum_{\substack{V_{-i} = x \\ v, v' \text{ differ on} \\ \text{place of winner} \\ \text{for } x}} \text{prob}(v_{-i} = x) \left(G(v, V_{-i}) - G(v', V_{-i})\right).$$

□

Lemma 2. *Consider $k = \#A \geq 3$. Suppose $f : \mathcal{L}^N \to A$ is such that for all $a \in A$ and for all voters $\mathcal{USR}_i$ there exists some choice of votes $V_{-i,const} \in \mathcal{L}^{N-1}$ for the other voters such that for all $v_i \in \mathcal{L}$, the resulting social choice is $f(v_i, V_{-i,const}) = a$. Consider two votes v_i and v_i' that differ only by a single transposition of adjacent elements: $v_i = [...]cd[...]$, $v_i' = [...]dc[...]$. Suppose that there exists a set of votes $V_{-i,var}$ for the other participants such that $f(v_i, V_{-i,var}) = a \neq f(v_i', V_{-i,var}) = b$. Then, if*

- *$a = d$ or*
- *$b = c$, $a \neq d$*

there exists a Schelling-dishonest evaluation where all votes available to $\mathcal{USR}_i$ are Schelling-dishonest.

Proof. Denote by n the number of options and j the position of the option d in v_i. Define $\epsilon = \frac{1}{6\left(\frac{n(n+1)}{2} - j + \frac{7}{2}\right)}$ and $z = 2\epsilon = \frac{1}{3\left(\frac{n(n+1)}{2} - j + \frac{7}{2}\right)}$.

Consider either of the cases where $a = d$ (with either $b = c$ or $b \neq c$ possible). Then consider a evaluation composed of:

- $V_{-i,var}$ with probability z.
- The rth choice of V_{-i} with constant probability $2r(z + \epsilon)$ for $r > j$.
- The rth choice of V_{-i} with constant probability $2(r-1)(z+\epsilon)$ for $r < j-1$.
- The choice c with constant probability $2j(z + \epsilon) + \epsilon$.
- The choice d with constant probability $2j(z + \epsilon)$.

Then note that

$$2(z+\epsilon)\left(\sum_{r=j+1}^{n} i + \sum_{r=1}^{j-2}(i-1) + 2j\right) + \epsilon + z = 2(z+\epsilon)\left(\frac{n(n+1)}{2} - j + \frac{7}{2}\right) = 1,$$

namely the probabilities of the different cases add to 1.

However, as $z = 2\epsilon > \epsilon$, we note in each case that the total probability of d winning (between the constant and variable V_{-i}) is greater than the total probability of c winning for v_i, and the total probability of c winning is greater than the total probability of d winning for the v_i' (whether $b = c$ or not). Hence, neither of these two votes is Schelling-honest. However, as all other pairs of options differ in their probability of winning by at least $2(z + \epsilon)$, and at most a probability of z can change between options in different votes, no other vote can be Schelling-honest.

For the case $b = c$, $a \neq d$, we employ the same argument using an evaluation with the same probabilities as above only where the probabilities of the constant $V_{-i,const}$ for c and d are reversed. Namely: the choice d with constant probability $2j(z + \epsilon) + \epsilon$ and the choice c with constant probability $2j(z + \epsilon)$.

□

Lemma 3. *Consider $k = \#A \geq 3$. Suppose $f : \mathcal{L}^N \to A$ is neutral and G is strictly reasonable. Further suppose that for some $w \in A$ and for all voters $\mathcal{USR}_i$ there exists some choice of votes $V_{-i,const} \in \mathcal{L}^{N-1}$ for the other voters such that for all $v_i \in \mathcal{L}$, the resulting social choice is $f(v_i, V_{-i,const}) = w$. Consider two votes v_i and v_i' that differ only by a single transposition of adjacent elements: $v_i = [...]cd[...]$, $v_i' = [...]dc[...]$. Suppose that there exists a set of votes $V_{-i,var}$ for the other participants such that $f(v_i, V_{-i,var}) = a \neq f(v_i', V_{-i,var}) = b$. Then, if*

- *$c \neq a, b$, $d \neq a, b$*
- *$d = b$, $c \neq a$, and $a >_{v_i} d$,*

there exists a Schelling-dishonest evaluation where $\mathcal{USR}_i$ can make both a Schelling-honest and a Schelling-dishonest vote, with the highest payoff available to her coming from a Schelling-dishonest vote.

Proof. Suppose we are in one of the cases:

- $c \neq a, b$, $d \neq a, b$, $a >_{v_i} b$ or
- $d = b$, $c \neq a$, $a >_{v_i} b$.

As f is neutral we can permutate $V_{-i,const}$ interchanging w with other options; hence we know that there exists some choice of the votes of other participants so that any given option wins regardless of the vote of $\mathcal{USR}_i$. In particular, we can assume without loss of generality that $w = d$. Take σ as the transposition that swaps c and d. Then take $V_{-i,const=c} = \sigma(V_{-i,const})$. Furthermore, as $\sigma(v_i) = v_i'$ and $\sigma(v_i') = v_i$, by the symmetry assumptions in Defintion 6, we have $G(v_i, V_{-i,const=c}) = G(v_i', V_{-i,const})$ and $G(v_i, V_{-i,const}) = G(v_i', V_{-i,const=c})$.

By strict reasonableness, as v_i' ranks d higher than v_i does, $G(v_i', V_{-i,const}) - G(v_i, V_{-i,const}) > 0$. Also, as we have assumed $a >_{v_i} b$ in these cases, by strict reasonableness $G(v_i, V_{-i,var}) - G(v_i', V_{-i,var}) > 0$.

Then, denote by n the number of options and j the position of the option d in v_i. Define

$$m = \max\left\{1, \frac{G(v_i', V_{-i,const}) - G(v_i, V_{-i,const})}{G(v_i, V_{-i,var}) - G(v_i', V_{-i,var})}\right\},$$

$$\epsilon = \frac{1}{(2m+1)(n(n+1) - 2j + 7)},$$

and

$$z = 2\epsilon m.$$

Note, in particular, $z > \epsilon$ and

$$z > \epsilon \frac{G(v_i', V_{-i,const}) - G(v_i, V_{-i,const})}{G(v_i, V_{-i,var}) - G(v_i', V_{-i,var})}.$$

Then consider an evaluation composed of:

- $V_{-i,var}$ with probability z.
- The rth choice of V_{-i} with constant probability $2r(z+\epsilon)$ for $r > j$.
- The rth choice of V_{-i} with constant probability $2(r-1)(z+\epsilon)$ for $r < j-1$.
- The choice c with constant probability $2j(z+\epsilon)$.
- The choice d with constant probability $2j(z+\epsilon) + \epsilon$.

Then note that

$$2(z+\epsilon)\left(\sum_{r=j+1}^{n} i + \sum_{r=1}^{j-2}(i-1) + 2j\right) + \epsilon + z = 2(z+\epsilon)\left(\frac{n(n+1)}{2} - j + \frac{7}{2}\right) = 1,$$

by our choices of ϵ and z. Namely the probabilities of the different cases add to 1. (Furthermore, as these probabilities are each clearly positive, they are all in $(0,1)$.)

As $\epsilon > 0$, the vote v_i' is Schelling-honest by construction. However, using Lemma 1 and noting that only the $j-1$ and jth places differ between v_i and v_i', we compute

$$E[G(v_i) - G(v_i')]$$

$$= z\left(G(v_i, V_{-i,var}) - G(v_i', V_{-i,var})\right) + \epsilon\left(G(v_i, V_{-i,const}) - G(v_i', V_{-i,const})\right) > 0,$$

by our choices of z and ϵ.

Hence, v_i has a higher expected payoff than v_i', but d has a greater total probability of winning than c when $\mathcal{USR}_i$ votes v_i. Thus, v_i is Schelling-dishonest. Moreover, as all other pairs of options differ in their probability of winning by at least $2(z+\epsilon)$, and at most a probability of z can change between options in different votes, no other vote can be Schelling-honest. So the vote with the highest payoff must be Schelling-dishonest.

The case where $c \neq a, b,\ d \neq a, b,\ a <_{v_i} b$ is similar where c has constant probability $2j(z+\epsilon)+\epsilon$, d has constant probability $2j(z+\epsilon)$, we take $m = \frac{G(v_i, V_{-i,const=c}) - G(v_i', V_{-i,const=c})}{G(v_i', V_{-i,var}) - G(v_i, V_{-i,var})}$, and v_i' is the Schelling-dishonest but higher payoff vote.

$\square$

Lemma 4. *Consider $k = \#A \geq 3$. Suppose $f : \mathcal{L}^N \to A$ is neutral and G is strictly reasonable. Further suppose that for some $w \in A$ and for all voters $\mathcal{USR}_i$ there exists some choice of votes $V_{-i,const} \in \mathcal{L}^{N-1}$ for the other voters such that for all $v_i \in \mathcal{L}$, the resulting social choice is $f(v_i, V_{-i,const}) = w$. Consider two votes v_i and v_i' that differ only by a single transposition of adjacent elements: $v_i = [...]cd[...]$, $v_i' = [...]dc[...]$. Suppose that there exists a set of votes $V_{-i,var}$ for the other participants such that $f(v_i, V_{-i,var}) = a \neq f(v_i', V_{-i,var}) = d$. Then, if $c \neq a$, and $a <_{v_i} d$, there exists a Schelling-dishonest evaluation.*

Proof. Denote by n the number of options and j the position of the option d in v_i. We consider two cases. First suppose that there exists some $v_i'' \in \mathcal{L}$ such that $f(v_i'', V_{-i,var}) = h$ is in the jth position of v_i''. Then v_i'' is of the form $[...]gh[...]$, with the option h being in the jth position. We consider the vote given by transposing g and h, $v_i''' : [...]hg[...]$. Then if $f(v_i''', V_{-i,var}) \neq h$, there must exist an evaluation where there are no possible Schelling-honest votes by (the first case of) Lemma 2. Hence, we can assume that $f(v_i''', V_{-i,var}) = h$.

Take σ to be the permutation of the options such that $\sigma(v_i'') = v_i'$. In particular, $\sigma(h) = c$ and $\sigma(g) = d$. Denote the kth element of a vote v by $v(k)$. Then, thinking of σ as a function $\sigma : A \to A$, we have that for $k \neq j-1, j$,

$$\sigma(v_i'''(k)) = \sigma(v_i''(k)) = v_i'(k) = v_i(k).$$

Hence, the same permutation σ also gives $\sigma(v_i''') = v_i$.

Then by neutrality, $f(\sigma(V_{-i,var}), \sigma(v_i'')) = \sigma(h) = c$ and $f(\sigma(V_{-i,var}), \sigma(v_i''')) = \sigma(h) = c$. We define $V_{-i,var2} = \sigma(V_{-i,var})$. Now we compute

$$G(v_i', V_{-i,var2}) = G(\sigma(v_i''), \sigma(V_{-i,var})) = G(v_i'', V_{-i,var}) > G(v_i, V_{-i,var}),$$

where we use the symmetry properties of reasonableness in Definition 6 for the second equality and we use strict reasonableness and the fact that $f(v_i'', V_{-i,var}) = v_i''(j)$, whereas $v_i(j) = d >_{v_i} a = f(v_i, V_{-i,var})$ for the inequality. Similarly, we see

$$G(v_i, V_{-i,var2}) = G(\sigma(v_i'''), \sigma(V_{-i,var})) = G(v_i''', V_{-i,var}) = G(v_i', V_{-i,var}),$$

where we again use both properties of reasonableness, this time with the fact that $v_i'''(j-1) = f(v_i''', V_{-i,var})$ and $v_i'(j-1) = f(v_i', V_{-i,var})$.

We now follow an argument similar to that of Lemma 3. By neutrality we can without loss of generality take $w = c$. Then take σ_2 to be the permutation that transposes c and d and take $V_{-i,const=d} = \sigma_2(V_{-i,const})$. Furthermore, as $\sigma_2(v_i) = v_i'$ and $\sigma_2(v_i') = v_i$, by the symmetry assumptions in Defintion 6, we have $G(v_i, V_{-i,const=d}) = G(v_i', V_{-i,const})$ and $G(v_i, V_{-i,const}) = G(v_i', V_{-i,const=d})$.

By strict reasonableness, as v_i ranks c higher than v_i' does, $G(v_i, V_{-i,const}) - G(v_i', V_{-i,const}) > 0$.

Define

$$m = 2 \max \left\{ 1, \frac{G(v_i, V_{-i,const}) - G(v_i', V_{-i,const})}{G(v_i', V_{-i,var2}) - G(v_i, V_{-i,var})} \right\},$$

$$\epsilon = \frac{1}{(2m+1)(n(n+1) - 2j + 7)},$$

and

$$z = 2\epsilon m.$$

Note, in particular, $z > 2\epsilon$ and

$$z > 2\epsilon \frac{G(v_i, V_{-i,const}) - G(v_i', V_{-i,const})}{G(v_i', V_{-i,var2}) - G(v_i, V_{-i,var})}.$$

Then consider an evaluation composed of:

- $V_{-i,var}$ with probability $\frac{z}{2}$.
- $V_{-i,var2}$ with probability $\frac{z}{2}$.
- The rth choice of v_{-i} with constant probability $2r(z + \epsilon)$ for $r > j$.
- The rth choice of v_{-i} with constant probability $2(r-1)(z+\epsilon)$ for $r < j-1$.
- The choice c with constant probability $2j(z + \epsilon) + \epsilon$.
- The choice d with constant probability $2j(z + \epsilon)$.

Then note that

$$2(z+\epsilon) \left(\sum_{r=j+1}^{n} i + \sum_{r=1}^{j-2} (i-1) + 2j \right) + \epsilon + z = 2(z+\epsilon) \left(\frac{n(n+1)}{2} - j + \frac{7}{2} \right) = 1,$$

by our choices of ϵ and z. Namely the probabilities of the different cases add to 1. (Furthermore, as these probabilities are each clearly positive, they are all in $(0, 1)$.)

We use Lemma 1 and note that only the $j-1$ and jth places differ between v_i and v_i'. Furthermore, we use the equality and inequalities found above on payoffs for different votes to compute

$$E[G(v_i') - G(v_i)] = \frac{z}{2} \left(G(v_i', V_{-i,var2}) - G(v_i, V_{-i,var}) \right) + \epsilon \left(G(v_i', V_{-i,const}) - G(v_i, V_{-i,const}) \right).$$

Then $E[G(v_i') - G(v_i)] > 0$ by our choices of z and ϵ.

Hence, v_i' has a higher expected payoff than v_i, but c has a greater total probability of winning than d when $\mathcal{USR}_i$ votes v_i'. Thus, v_i' is Schelling-dishonest. Moreover, as all other pairs of options differ in their probability of winning by at least $2(z + \epsilon)$, and at most a probability of z can change between options in different votes, no other vote can be Schelling-honest. So the vote with the highest payoff must be Schelling-dishonest in this case.

Now suppose that there does not exist any $v_i'' \in \mathcal{L}$ such that $f(v_i'', V_{-i,var})$ is in the jth position of v_i''. As $a <_{v_i} d$, v_i has the form $[...]cd[...a...]$. Consider the transposition of v_i that moves a up one place. Then if the resulting vote does not go to a we have a pair of votes that correspond to (the first case of) Lemma 2, implying the existence of a dishonest evaluation. However, as long as this process does not produce a Schelling-dishonest evaluation, it can be continued until eventually we transpose a with d to produce a vote of the form $v_i''' : [...]cad[...]$. As a is then in the jth position of this vote, by assumption, $f(v_i''', V_{-i,var}) \neq a$. Then, finally, we must have a Schelling-dishonest evaluation by applying Lemma 2 to this last transposition.

□

We now have the tools to prove the above result.

Proof of Theorem 1. As we assume that f has AV, it is also Pareto efficient and non-dictatorial. Hence by Theorem A of [20], we know that f is not (strongly) monotonic. (This strong sense of monotonicity used in [20], which differs from that of Definition 3, includes properties related to independence of irrelevant alternatives. Here we are essentially applying a version of the Gibbard-Satterthwaite Theorem.)

Specifically, there exist $l_1, \ldots, l_N$ such that $f(l_1, \ldots, l_N) = a$, $l_1', \ldots, l_N'$ such that for all b each l_i' ranks $a > b$ if l_i does, but $f(l_1', \ldots, l_N') \neq a$.

Then we have a sequence

$$a = f(l_1, \ldots l_N), f(l_1, \ldots l_{N-1}, l_N'), \ldots, f(l_1, l_2' \ldots, l_N'), f(l_1', \ldots, l_N') \neq a$$

formed by replacing the l_j by the l_j' one at a time. Take i to the smallest value at which this replacement results in a choice choice other than a.

Note that as $a >_{l_i'} b$ for all b such that $a >_{l_i} b$, there exists a series of elements in $\mathcal{L}$: $t_1, \ldots t_r$ from l_i to l_i' each differing by a transposition of adjacent elements σ_j such that none of the σ_j move a down.

Similarly to above, as

$$f(l_1, \ldots, l_{i-1}, l_i, l_{i+1}', \ldots, l_N') = a$$

and

$$f(l_1, \ldots, l_{i-1}, l_i', l_{i+1}', \ldots, l_N') \neq a,$$

we can think of $\mathcal{USR}_i$'s vote being changed one transposition at a time and take k as the smallest value such that

$$f(l_1, \ldots, l_{i-1}, t_k, l_{i+1}', \ldots, l_N') \neq a.$$

Hence, we have $l_1, \ldots, l_{i-1}, l'_{i+1}, \ldots, l'_N$, t_{k-1}, and t_k,

- where t_{k-1} and t_k differ by a transposition σ_k of adjacent elements that either moves a up or fixes a, and changes the result from a to b
- $f(l_1, \ldots, l_{i-1}, t_{k-1}, l'_{i+1}, \ldots, l'_N) = a$
- $f(l_1, \ldots, l_{i-1}, t_k, l'_{i+1}, \ldots, l'_N) = b \neq a$

Denote the two elements being transposed by σ_k as c and d, so that t_{k-1} has the form [...]cd[...] and t_k has the form [...]dc[...] (where as above terms in brackets are fixed between the two orders). Then we have several cases:

- $a = d$ OR $b = c$, $a \neq d$. Here we construct an evaluation where there is no Schelling-honest vote using Lemma 2,
- $c \neq a, b$, $d \neq a, b$, OR $b = d$, $a \neq c$, $a >_{v_i} b$. Here we construct an evaluation where there is a Schelling-honest vote, but the highest payoff comes from a Schelling-dishonest vote using Lemma 3.
- $b = d$, $a \neq c$, $a <_{v_i} d$. Here we construct a Schelling-dishonest evaluation (in which a Schelling-honest vote may or may not exist) using Lemma 4.

Hence, in each case there exists some Schelling-dishonest evaluation. (Note that any cases where $a = c$ are already excluded as we know that σ_k does not move a down.)

□

8 Examples of incentive compatible systems failing the criteria of Theorem 1

In this section we see counterexamples to version of Theorem 1 when one attempts to relax some of the assumptions. First we see an example where G is only assumed to be reasonable rather than strictly reasonable.

Example 1. *Suppose participants rank all possible outcomes creating an ordering in $\mathcal{L}$ and place a deposit D. The winning outcome is decided by Plurality voting and payoffs are given by the the "all or nothing" payoff system of Definition 7.*

Then for a given $v = a_1 > a_2 > \ldots$, the user's expected return is

$$-D \cdot (1 - prob(a_1 \ wins | \mathcal{USR}_i \ votes \ v)) = -D \cdot (1 - p(a_1)),$$

using the notation of Section 5.

Note that this is optimized by maximizing $prob(a_1 \ wins : \mathcal{USR}_i \ votes \ v)$. We claim that it is always in the voter's interest to produce a vote v such that

$$prob(a_1 \ wins : \mathcal{USR}_i \ votes \ v) \geq prob(a_j \ wins : \mathcal{USR}_i \ votes \ v) \qquad (1)$$

for all a_j. If this were not the case, then by voting $v_2 = a_j > a_1 > a_2 > \ldots > a_{j-1} > a_{j+1} > \ldots$, the voter would have

$$prob(a_j\ wins : \mathcal{USR}\ votes\ v_2) \geq prob(a_j\ wins : \mathcal{USR}\ votes\ v) > prob(a_1\ wins : \mathcal{USR}\ votes\ v)$$

as Plurality is a monotonic voting system (for any given possibility in V_{-i} of how the other voters vote, ranking a_j higher cannot cause it to lose when $\mathcal{USR}_i$ votes v_2 scenarios it would have won if $\mathcal{USR}_i$ votes v). Hence, whichever a_j that produces the largest value of $prob(a_j\ wins$: modified v to move a_j first$)$, both maximizes the payoff and satisfies inequality 1.

Then, as according to the payoff scheme of Definition 7 only the first choice vote of a user has any effect on either the winner or the payout, $\mathcal{USR}_i$ does no worse by choosing $a_2, a_3, \ldots$ such that

$$prob(a_j\ wins : \mathcal{USR}\ votes\ v) \geq prob(a_{j+1}\ wins : \mathcal{USR}\ votes\ v)$$

for all j.

Now we consider an example where the social choice function is Pareto efficient and non-dictatorial, but fails neutrality and absence of vetos. This example can be performed with any (strictly reasonable) payoff function.

Example 2. *Consider the following voting system when there are three options $\{a, b, c\}$. If the first choices of voters are unanimous, their unanimous first choice is selected. Otherwise, the outcome is a. This system is clearly Pareto efficient and non-dictatorial, but it does not satisfy neutrality or absence of veto. Then, under a strictly reasonable payoff system, voting $a > b > c$ (or voting $a > c > b$ which by reasonableness has the same payoff) is always a Schelling-honest vote that maximes a user's payoff.*

Note that as Example 2 gives a non-dictatorial game with three outcomes that is incentive compatible in the sense of Section 5, we have that Schelling-honesty cannot be made to fit into the notion of "straightforwardness" as expressed in [12], even though that notion is very general. Indeed, in the context of Schelling games the honest vote of a participant *depends* on the user's beliefs about the votes of others, which is not covered in the framework of straightforwardness.

9 Existence of honest votes

In Section 7 we saw that, when certain natural requirements are placed on a Schelling game voting and payoff system, there will inevitably be situations where it is in the economic interest of participants to cast Schelling-dishonest votes in the sense of Definition 4. As part of that argument, in Lemma 2, we considered cases where no Schelling-honest vote is available. However, in the proof of Theorem 1, it is not shown that the situation of Lemma 2 is ever necessarily realized. In this section we will consider explicit situations where it is impossible for a voter to cast a Schelling-honest vote in several specific voting systems.

Example 3. *Consider an Instant-runoff voting system that is attempting to decide between three outcomes* $\{a, b, c\}$.

Suppose $\mathcal{USR}_i$ *has two votes (this avoids having situations with ties). Before voting, her probability function for the voters of others* V_{-i} *is given as follows*

- *There is a* 49% *chance that there will be*
  - 21 *votes for* $a > b > c$
  - 22 *votes for* $b > c > a$
  - 20 *votes for* $c > a > b$
- *There is a* 49% *chance that there will be*
  - 20 *votes for* $a > b > c$
  - 21 *votes for* $b > c > a$
  - 22 *votes for* $c > a > b$
- *There is a* 2% *chance that there will be*
  - 22 *votes for* $a > b > c$
  - 20 *votes for* $b > c > a$
  - 21 *votes for* $c > a > b$

Then, regardless of which of the possibilities in $\mathcal{L}^{N-1}$ *occurs and regardless of how* $\mathcal{USR}_i$ *votes,* a *wins any eventual duels against* b, b *wins any eventual duels against* c, *and* c *wins any eventual duels against* a.

If $\mathcal{USR}_i$ *ranks* a *first with her two votes, then there is a* 49% *chance that the runoff will be between* a *and* b *and a* 51% *chance that the runoff will be between* a *and* c. *Hence* a *has a* 49% *chance of winning,* b *a* 0% *chance of winning, and* c *a* 51% *chance of winning.*

If $\mathcal{USR}_i$ *ranks* b *or* c *first with her two votes we have similar results:*

$\mathcal{USR}_i$ \vote outcome	a	b	c
a *first*	49%	0%	51%
b *first*	51%	49%	0%
c *first*	0%	98%	2%

Hence, in any case, $\mathcal{USR}_i$ *does not produce a vote in which the first choice is the outcome with the highest chance of winning. Then, for a user in this situation, the strategy of producing a list ordered by probability cannot be incentive compatible because it is impossible to follow. This is true regardless of the system of rewards and punishments applied.*

Further note that this example also applies to the Smith-IRV system, namely the system of performing IRV but checking before each round whether there is a Condorcet winner and selecting this outcome to be the winner if there is.

We already saw in Example 1 that in some systems it is possible to always have a Schelling-honest vote. We might ask if this property is compatible with other properties of voting systems, such as the Condorcet criterion.

Example 4. *Consider the voting system that*

- *Checks to see if there is a Condorcet winner, is so this choice is declared the winner*
- *If there is no Condorcet winner, the Plurality winner is selected.*

One might hope that this would be a reasonable candidate for a system that would always have a Schelling-honest vote considering its relation to the Plurality voting system and 1. However, this is not the case. Consider the situation where $\mathcal{USR}$ *has three votes and has an evaluation of*

- *Scenario 1: There is a* 34% *chance that*
  - 24 *votes for* $b > c > a$
  - 22 *votes for* $b > a > c$
  - 27 *votes for* $a > b > c$
  - 27 *votes for* $c > a > b$

$\mathcal{USR}$ vote	$a > b > c$	$a > c > b$	$b > a > c$	$b > c > a$	$c > a > b$	$c > b > a$
outcome	a	a	a	b	b	b

- *Scenario 2: There is a* 33% *chance that one has the appropriate permutation of Scenario 1 so that we have*

$\mathcal{USR}$ vote	$a > b > c$	$a > c > b$	$b > a > c$	$b > c > a$	$c > a > b$	$c > b > a$
outcome	a	c	a	a	c	c

- *Scenario 3: There is a* 33% *chance that one has the appropriate permutation of Scenario 1 so that we have*

$\mathcal{USR}$ vote	$a > b > c$	$a > c > b$	$b > a > c$	$b > c > a$	$c > a > b$	$c > b > a$
outcome	c	c	b	b	c	b

This leads to percent chances of winning of

$\mathcal{USR}_i$ \vote outcome	a	b	c
$a > b > c$	67%	0%	33%
$a > c > b$	34%	0%	66%
$b > a > c$	67%	33%	0%
$b > c > a$	33%	67%	0%
$c > a > b$	0%	34%	66%
$c > b > a$	0%	67%	33%

Here there are votes where the first choice has the highest chance of winning, but there is no vote where all options are ranked in order by chance of winning.

Proposition 2. *Suppose we have a monotonic voting system. Then whatever a user's evaluation V_{-i} about the votes she does not control, there exists a vote v_j that she can cast so that the highest ranked choice of v_j is the outcome with the highest probability of winning.*

Proof. Note that there are a finite number of permutations of the outcomes. Each permutation, if submitted as a vote, yields a percentage of each of its outcomes winning. Take the permutation

$$v_1 : a_{\sigma(1)} > \ldots > a_{\sigma(n)}$$

that maximizes the probability that its first choice wins. A priori, it is possible that some lower ranked choice $a_{\sigma(j)}$ has an even higher probability of winning, $p(a_{\sigma(j)}) > p(a_{\sigma(1)})$. Then consider the situation where the user votes

$$v_2 : a_{\sigma(j)} > a_{\sigma(1)} > \ldots > a_{\sigma(j-1)} > a_{\sigma(j+1)} > \ldots > a_{\sigma(n)}.$$

By monotonicity, for any collection of votes in $\mathcal{L}^{N-1}$ in which $a_{\sigma(j)}$ wins if $\mathcal{USR}_i$ votes v_1, $a_{\sigma(j)}$ still wins if $\mathcal{USR}_i$ votes v_2. Hence v_2 gives a vote for which the first choice has a probability of winning greater than $p(a_{\sigma(1)})$, contradicting the assumed maximality of $a_{\sigma(1)}$. □

Example 5. *We see that even with a monotonic voting system, it is still possible to have situations where there is no Schelling-honest vote. Suppose that $\mathcal{USR}_i$ controls three votes, and her probability distribution for the other voters is given as described below. Then, note that in each of the Ranked Pairs, Minmax, Kemeny-Young, and Schulze systems, in each scenario the outcome is given as in the tables below. However, these systems are Condorcet and monotonic, (see for example [27] for a discussion of why Ranked Pairs possesses these properties).*

- *Scenario 1: There is a 10% chance that there will be*
  - 14 *votes for* $a > b > c$
  - 3 *votes for* $a > c > b$
  - 15 *votes for* $b > c > a$
  - 16 *votes for* $c > a > b$

$\mathcal{USR}_i$ *vote*	$a > b > c$	$a > c > b$	$b > a > c$	$b > c > a$	$c > a > b$	$c > b > a$
outcome	a	c	a	c	c	c

- *Scenario 2: There is a 32% chance that there will be*
  - 12 *votes for* $a > b > c$
  - 3 *votes for* $a > c > b$
  - 17 *votes for* $b > c > a$
  - 16 *votes for* $c > a > b$

$\mathcal{USR}_i$ *vote*	$a>b>c$	$a>c>b$	$b>a>c$	$b>c>a$	$c>a>b$	$c>b>a$
outcome	c	c	b	b	c	c

- *Scenario 3: There is a 17% chance that one is in a scenario where b wins regardless of the vote of* $\mathcal{USR}_i$*, such as*
  - *48 votes for* $b>a>c$
- *Scenario 4: There is a 26% chance that there will be*
  - *13 votes for* $a>b>c$
  - *6 votes for* $a>c>b$
  - *18 votes for* $b>c>a$
  - *11 votes for* $c>a>b$

$\mathcal{USR}_i$ *vote*	$a>b>c$	$a>c>b$	$b>a>c$	$b>c>a$	$c>a>b$	$c>b>a$
outcome	a	a	a	b	c	b

- *Scenario 5: There is a 15% chance that one is in a scenario where a wins regardless of the vote of* $\mathcal{USR}_i$*, such as*
  - *48 votes for* $a>b>c$

Then, we summarize which scenarios and which user votes give each outcome:

$\mathcal{USR}_i$ *\vote outcome*	a	b	c
$a>b>c$	*1,4,5*	*3*	*2*
$a>c>b$	*4,5*	*3*	*1,2*
$b>a>c$	*1,4,5*	*2,3*	
$b>c>a$	*5*	*2,3,4*	*1*
$c>a>b$	*5*	*3*	*1,2,4*
$c>b>a$	*5*	*3,4*	*1,2*

In percent chance of winning:

$\mathcal{USR}_i$ *\vote outcome*	a	b	c
$a>b>c$	51%	17%	32%
$a>c>b$	41%	17%	42%
$b>a>c$	51%	49%	0%
$b>c>a$	15%	75%	10%
$c>a>b$	15%	17%	68%
$c>b>a$	15%	43%	42%

Remark 1. *Consider a monotonic voting system and suppose one begins with* $a>b>c$ *as the vote guaranteed by Proposition 2 that places the highest ranked choice first. Then one could attempt to obtain a Schelling-honest vote by permuting two options at a time: e.g. if* $a>b>c$ *yields* $p(c)>p(b)$ *then try*

$a > c > b$. By monotonicity, it is still the case that $p(c) > p(b)$, but it is possible that $p(c) > p(a)$. If this is the case one can try $c > a > b$. By monotonicity, it is then the case that $p(c) > p(a)$, but it is possible that $p(b) > p(a)$, etc. Continuing like this any evaluation that $\mathcal{USR}_i$ might have that results in no possible Schelling-honest vote in a monotonic voting system (at least with three outcomes) must have a particular form. Then, we see in Example 5 that that form can be realized in the Ranked Pairs, Minmax, Kemeny-Young, and Schulze systems.

Given these examples, we ask:

Question 1. *Does there exist any Condorcet voting system in which it is always possible to provide a Schelling-honest vote, i.e. a ranking in order by the probability that each outcome wins, once your vote is taken into account and for an arbitrary evaluation of the distribution of the other votes?*

10 Conclusion and future work

In this work we have begun to look at a version of social choice theory in the context where participants in an election are motivated by the economic incentives of a Schelling game rather than by their own candidate preferences. We have provided a definition for what "honesty" can mean in this framework, and we have observed that many of the subtleties of traditional social choice theory regarding tactical voting have analogs here.

In particular, we saw an impossibility result in Theorem 1 that shows that some of the properties that one would hope an ideal Schelling game based voting system to have are incompatible. Ultimately this result is similar to Arrow's Impossibility Theorem and the Gibbard-Satterthwaite Theorem in that they only necessarily apply to a fairly narrow range of potential voting scenarios. With many voting systems, honest behaviour is incentivized most of the time. Moreover, just as the classic impossibility theorems of social choice theory do not render standard elections unusable for determining collective decisions, Schelling game based voting is still likely to be useful as a tool in applications to blockchain oracles, and perhaps beyond to other crowdsourcing platforms.

Thus, in the absence of ideal voting systems, one is left with the task of considering which partial properties one can establish for such systems and the potential tradeoffs between goals that are potentially incompatible (e.g. such as the incompatibility between the Condorcet criteria and escaping the no-show paradox, [17]). We conclude by considering more broadly some the properties that one would want Schelling game voting systems to have.

- We want voters to be incentivized as much as possible to *actually* provide an "honest" list of choices. For example, we would want
  - Some partial guarantees[2] limiting the failures of incentive compatibility for Schelling-honesty.

[2]One potential such approach is the following: In some Schelling game based blockchain

- Clone independence. Note that while options may have identical implications when implemented (on a blockchain platform or otherwise) leading them to be reasonably thought of as clones, as the voters themselves do not have intrinsic preferences the standard definition of clones [27] no longer makes sense. Thus, notions of clone independence are not included in our notion of Schelling-honesty, in contrast to their inclusion in non-manipuability in standard elections. Nonetheless, to the degree that a collection of clones reduces the chances of any given clone to win, their presence and how they are treated by a voting system may have implications for parties interested in the outcome of a given vote. Note that these issues touch on the concept of "control", see Chapter 7 of [5].
- Attack resistance based on the idea that it should be expensive to bribe sufficient numbers of voters to change the outcome of a decision, at least relative to the values that can be gained by successful attacks on a given crowdsourcing platform. Here as voters are assumed to be incentivized purely economically, analyzing the bribes required for them to change their votes is tractable. Also, one can clearly apply work in computational social choice theory that estimates the computational difficulty of finding which sets of voters are sufficient to be bribed, see Section 7 of [5]. Note that the framework of Schelling games opens up new possibilities for game-theoretic attacks [6] that must be considered.

• Then we want the lists voters provide to be converted into a collective outcome in such a way that if voters provide "honest" lists, the collective choice will also be "honest". For this we might want,

  - A voting rule that is a Maximum Likelihood Estimator under some reasonable noise model, following the work of [8]. Indeed, in the ideal where the above considerations manage to incentivize participants to submit votes that reflect the answer they observe, one is back in the framework of the kind of crowdsourced decision making considered by [8].

• Finally, use on a blockchain platform may impose additional constraints such as

  - New running time considerations. Already in standard voting systems, there are interesting questions related to the running time of

oracle applications there is a notion of appeal, where payoffs are determined based on an ultimate outcome decided in a later round. In such situations, a voter $\mathcal{USR}_i$ that expected an appeal might belief the outcome of such a process to be independent of her vote. Under this (rather strong) assumption, one can see that a wider range of voting and payoff systems can be seen to incentivize Schelling-honesty in the presence of appeals. Compare Theorem 1 to Proposition 1 of [15].

tabulating votes, with implications on the viability of different voting systems, see for example discussions related to "winner problem" complexity for various voting system in [5]. However, in the domain of blockchains, as on-chain computation is performed by a large number of verifiers and consequently has elevated cost, it is often advantageous to minimized such on-chain computation by using processes that allow for parties to submit an answer, computed off-chain, that can be efficiently *verified* on-chain, in the style of [26]. Then one asks for which voting rules can vote tabulation be effectively adapted to such a framework.

Then, future work considering to what degree these goals are compatible seems fruitful.

References

[1] John Adler, Ryan Berryhill, Andreas G. Veneris, Zissis Poulos, Neil Veira, and Anastasia Kastania. ASTRAEA: A decentralized blockchain oracle. *2018 IEEE Confs on Internet of Things, Green Computing and Communications, Cyber, Physical and Social Computing, Smart Data, Blockchain, Computer and Information Technology, Congress on Cybermatics*, 2018.

[2] Douglas Ahler, Carolyn Roush, and Gaurav Sood. The micro-task market for lemons: Data quality on Amazon's mechanical turk. `http://www.gsood.com/research/papers/turk.pdf`. 2018.

[3] Kenneth J. Arrow. A difficulty in the concept of social welfare. *Journal of Political Economy*, 58(4):328–346, 1950.

[4] David Austen-Smith and Jeffrey S. Banks. Information aggregation, rationality, and the condorcet jury theorem. *American Political Science Review*, 90(1):34–45, 1996.

[5] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia. *Handbook of Computational Social Choice*. Cambridge University Press, New York, NY, USA, 1st edition, 2016.

[6] Vitalik Buterin. The p+epsilon attack. `https://blog.ethereum.org/2015/01/28/p-epsilon-attack/`. January 2015.

[7] Ioannis Caragiannis, Ariel D. Procaccia, and Nisarg Shah. When do noisy votes reveal the truth? In *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, EC '13, pages 143–160, New York, NY, USA, 2013. Association for Computing Machinery.

[8] Vincent Conitzer and Tuomas Sandholm. Common voting rules as maximum likelihood estimators. *CoRR*, abs/1207.1368, 2012.

[9] Russell Cooper. *Coordination Games*. Cambridge University Press, 1999.

[10] Nicolas de Condorcet. Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. `https://gallica.bnf.fr/ark:/12148/bpt6k417181/f6.image`. 1785.

[11] William George and Clément Lesaege. A Smart Contract Oracle for Approximating Real-World, Real Number Values. In *International Conference on Blockchain Economics, Security and Protocols (Tokenomics 2019)*, volume 71 of *OpenAccess Series in Informatics (OASIcs)*, pages 6:1–6:15, Dagstuhl, Germany, 2019. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[12] Allan Gibbard. Manipulation of Voting Schemes: A General Result. *Econometrica*, 41(4):587–601, July 1973.

[13] Patrick Hummel. Jury theorems with multiple alternatives. *Social Choice and Welfare*, 34(1):65–103, January 2010.

[14] Ryan Kennedy, Scott Clifford, Tyler Burleigh, Ryan Jewell, and Philip Waggoner. The shape of and solutions to the mturk quality crisis. *SSRN Electronic Journal*, 01 2018.

[15] Clément Lesaege, William George, and Federico Ast. Kleros: Long paper v1.0.0. `https://kleros.io/whitepaper_long_en.pdf`. March 2020.

[16] Andrew Mao, Ariel D. Procaccia, and Yiling Chen. Better human computation through principled voting. In *Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence*, AAAI13, pages 1142–1148. AAAI Press, 2013.

[17] Hervé Moulin. Condorcet's principle implies the no show paradox. *Journal of Economic Theory*, 45(1):53 – 64, 1988.

[18] Jack Peterson and Joseph Krug. Augur: a decentralized, open-source platform for prediction markets. `http://bravenewcoin.com/assets/Whitepapers/Augur-A-Decentralized-Open-Source-Platform-for-Prediction-Markets.pdf/`, 2015.

[19] Marcus Pivato. Voting rules as statistical estimators. *Social Choice and Welfare*, pages 581–630, 2011.

[20] Philip J. Reny. Arrow's theorem and the Gibbard-Satterthwaite theorem: a unified approach. *Economics Letters*, 70(1):99–105, January 2001.

[21] 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.

[22] T.C. Schelling. *The Strategy of Conflict*. Harvard University Press, 1980.

[23] W.C. Stirling. *Theory of Social Choice on Networks: Preference, Aggregation, and Coordination*. Cambridge University Press, 2016.

[24] Lars-Gunnar Svensson. The proof of the Gibbard-Satterthwaite theorem revisited, 1999.

[25] Paul Sztorc. Truthcoin: Peer-to-peer oracle system and prediction marketplace. `https://www.truthcoin.info/papers/truthcoin-whitepaper.pdf`. Version 1.5, December 2015.

[26] Jason Teutsch and Christian Reitwießner. A scalable verification solution for blockchains. `http://people.cs.uchicago.edu/~teutsch/papers/truebit.pdf`, 2017.

[27] T. N. Tideman. Independence of clones as a criterion for voting rules. *Social Choice and Welfare*, 4(3):185–206, Sep 1987.

[28] Vitalik Buterin. SchellingCoin: A minimal-trust universal data feed. Ethereum blog: `https://blog.ethereum.org/2014/03/28/schellingcoin-a-minimal-trust-universal-data-feed/`. March 2014.