Robustness of Voting Mechanisms to External Information

Yiling Chen[1] and Jessie Finocchiaro[2⋆][0000−0002−3222−0089]

[1] Harvard University, yiling@seas.harvard.edu
[2] Boston College, finocch@bc.edu

Abstract. Mechanism design and social choice critically rely on having access to preferences over outcomes from a set of participants, implicitly assuming these preferences are *known* and *correct*. However, using the example of voting, it is well-understood that voters often submit preferences that are dissonant with their policy positions [20, 29]. We propose a stylized model of representing informational externalities causing some of this dissonance, in which voters update and move their "true" (but unacted upon) preferences based on a piece of global, external information. We propose a modified voting rule that recovers true preferences while querying only ordinal, rather than cardinal preferences. This enables us to learn enough about the strength of preferences to understand changes to votes, without fully learning cardinal preferences. With this modified voting rule, we are then able to understand the robustness of a large class of voting rules to these informational externalities, bounding the probability of voting outcomes changing because of informational externalities.

1 Introduction

Classical results in voting theory prove that no voting mechanism, which is both non-dictatorial and anonymous, can be entirely immune to strategic manipulation [18, 30]. This phenomenon is often evident in real-world scenarios: for instance, a voter who supports candidate c might opt to vote for candidate a instead if candidates a and b are leading in the polls (e.g., [12]). This strategic behavior often appears to be "directed" in practice, influencing voters to align their preferences with a piece of external information. We explore how to learn a voter's true preferences, before *informational externalities*, without having to elicit the voter's full cardinal preferences which are often imprecisely known. When informational externalities are in line with societal preferences, we additionally examine the impact of these externalities on total utility over the population, measured in social welfare.

Understanding the effects of informational externalities on voting outcomes is not straightforward: most voting mechanisms only elicit *ordinal* preferences from voters, but *cardinal* preferences conveying the strength of preferences are typically needed to know when a voter will act differently when presented with externalities. For intuition, consider an externality such as the "community cardinal utility" representing the community's preferences, as the aggregation of a population's ordinal preferences through pre-polling or citizens' assemblies.

⋆ Corresponding author; finocch@bc.edu

Example 1. A news station is covering an election between candidates a, b, and c. They report that candidate a is estimated to receive 41% percent of the votes, candidate b is estimated at 43%, and candidate c is to receive 16%. If the voters who favor candidate c much prefer candidate a to b (e.g., $c \succ a \gg b$), they might strategically change their votes to put a as their most preferred candidate upon seeing these polls.

The reason for this third group of voters changing their votes is not our interest: voters might be benevolent, seeking outcomes that benefit the common good, or their deviation might be self-motivated. Regardless of the motivation, behavioral shifts like those in Example 1 are common. While Example 1 has a group of voters with preferences "$c \succ a \gg b$", cardinal utilities are necessary to quantify the exact strength of these preferences relative to each other. In practice, however, these cardinal preferences are difficult to specify and impractical to elicit. Without access to cardinal utilities, it is difficult to quantify the effect of informational externalities because we do not know if preferences are actually weak enough to actually change an individual's vote. Therefore, we must find some way to understand the effect of these externalities without fully eliciting cardinal utilities. Only once we have done this can we begin to study the impact of external information on voting outcomes.

To this end, we introduce an intermediate, asymmetric, voting rule in § 3 that enables us to compute how a voter would have voted in the absence of the informational externality. Now, with two queries to voters (one with the standard voting rule and one with the asymmetric voting rule), we can analyze how the externality changes voting outcomes in expectation *without needing access to cardinal utilities*. While we study average-case voting, our approach is quite different than previous works; instead of examining expected social welfare conditioned on a voting profile, we are interested in changes to expected social welfare under mild, but directed, perturbations to voters' cardinal preferences.

We view our two-query voting mechanism as our primary contribution (§ 3), and expect its guiding principles can be generalized to other types of mechanisms where participants submit ordinal preferences over alternatives (without money). Our intermediate voting rule further allows us to answer the following three questions:

1. In § 4.1, we obtain bounds on the probability an alternative is selected by a voting rule; how do these bounds change when voters are affected by informational externalities? (Corollaries 3 and 4)
2. When does the expected social welfare increase with informational externalities? (Theorem 2)
3. Can we bound the probability that social welfare decreases in the presence of informational externalities on any given voting instance? (Theorem 3)

The answers to these questions are quite intuitive: in question (1), we find that the bounds obtained in § 4.1 via a balls-to-bins approach to deriving bounds are indeed tightened *for the outcomes most favored by external information* when voters are affected by externalities. We find that Question (2) unfolds similarly: if the worst (from a social-welfare perspective) alternative that the externality promotes provides more social welfare than any alternative demoted, we can expect social welfare to increase. Finally, we answer Question (3) positively by using Chernoff bounds and properties of the moment generating function to provide a bound defined by the expected change in social welfare.

While these results themselves may be unsurprising, we find the methodology to establish these results generalizable, opening up avenues for future research into the informational robustness of commonplace mechanisms. In this work, we think of robustness of mechanisms not against some adversary or some uniformly random noise, but instead to some powerful, pervasive information source. In particular, one could think of the source of the informational externality as the dissemination of a citizens' assembly [27], or perhaps less benevolently, the influence of media monopoly that has control over public dissemination of information. We use voting as a case study in this paper to demonstrate the larger ideal of informational robustness, and proceed by examining some commonplace voting mechanisms (Plurality, Borda, and Veto).

1.1 Related Work

Informational externalities affecting preferences Recently, Flanigan *et al.* [17] use implicit utilities (term from [13]) to examine the distortion that emerges under *public-spirited voting*. Their model of public-spirited utilities is a direct analog of the model presented by Chen *et al.* [15], which studies the price of anarchy in congestion games with altruistic players. Our model of social voting is subtly different, but similar to, the previous models. Namely, Flanigan *et al.* [17] and Chen *et al.* [15], inspired by Ledyard [23, p. 154], model players and voters as moving towards the *optimal* social outcome, while we model voters that move towards *a given* social outcome. Their models have slightly more flexiblility on the weight of social outcomes as they analyze worst-case (distortion, price of anarchy) results in their respective settings, while we focus on *average-case* outcomes, which requires the additional assumption that all voters weigh the external information equally.[3] Bedaywi *et al.* [7] use the same model of voting as Flanigan *et al.* [17] to bound the distortion of participatory budgeting mechanisms with public-spirited voters. Finally, Boehmer *et al.* [10] formalized a model of "biased" ordinal preferences in the context of multiwinner voting.

Average-case voting While the majority of the computational social choice literature focuses on worst-case impossibilities and tradeoffs, a solid line of work has examined average-case voting outcomes. These works tend to examine voting outcomes in expectation over votes drawn i.i.d. from some density μ. When μ is the uniform density, this yields the *impartial culture* assumption, which is notably unrealistic, but an important starting point for understanding average-case voting [24, 34]. Boutilier *et al.* [11] and Apesteguia *et al.* [2] both study the expected social welfare in average-case voting under slightly more general assumptions than impartial culture. Both show that, if the impartial culture assumption is satisfied, then Borda voting maximizes the expected social welfare conditioned on the fixed ordinal votes. We re-iterate that our work differs from this literature as we examine changes to expected social welfare under mild, but directed, perturbations to voters' cardinal preferences. Our voting model is also similar to, but subtly

[3] Average-case voting analyses are based on the premise that votes are i.i.d. samples from some distribution; changing the weight voters put on the externality induces some unique distribution shift per voter, and therefore samples are no longer identically distributed.

different from metric voting, in which voters and *alternatives* are embedded into $\mathbb{R}^d$, then preferences are computed by evaluating relative distances to alternatives. Distortion bounds in metric voting models are well understood [1, 14, 33], and recently examined with informational externalities when these externalities are precisely the "optimal social welfare" [5]. In our model, instead of voters and *alternatives* being embedded into $\mathbb{R}^d$, we assume voters and *scores* are embedded into $\mathbb{R}^d$, and voters submit the *score* they are closest to in distance.

Robustness in social choice and mechanism design Robustness within the field of social choice (in particular, within voting) has been measured in a few ways, but usually examined within the scope of uniformly random noise, Mallows noise, or adversarial perturbations. For example, Elkind *et al.* [16] introduces *swap bribery*, which Shiryaev *et al.* [32] leverage as a measure of (adversarial) robustness. In contrast, Boehmer *et al.* [8, 9] takes an average-case view to evaluate voting outcomes, using *uniformly random noise* and Baumeister and Hogrebe [6] examines robustness of voting rules under Mallows noise and when the belief distribution μ is explicitly given. Weyl [35] additionally examines the robustness of the quadratic voting rule, deviating from the other literature through its implicit use of money in voting. Finally, Givi *et al.* [19] empirically study informational robustness in voting through different preference formation models, but do not give theoretical insights in their findings around likelihood of voting outcomes.

2 Voting Model

Suppose n voters are working together to select an alternative in $[m] := \{1, 2, \ldots, m\}$. Voter i has some normalized utility $u_i \in \Delta_m$ representing their (possibly unknown) *cardinal preferences*, where $\Delta_m = \{u \in \mathbb{R}^m_+ : \|u\|_1 = 1\}$ denotes the probability simplex over m alternatives. We suppose utilities are drawn from some unknown density μ defining the probability space $(\Delta_m, [m], \mu)$. Each voter must select some score vector r in a given *scoring menu* $\mathcal{R}$ representing *ordinal preferences* to assign to alternatives such that $|\mathcal{R}| = R < \infty$, where $\mathcal{R} \subset C\Delta_m$ for some $C > 0$ is a subset of a general simplex. [4] While the true underlying cardinal preferences u are often impractical to elicit, our voting rules instead elicit the report $r \in \mathcal{R}$ "corresponding to" u.

A set of votes is then input to the voting rule $f : \mathbb{N}^R \to [m]$ which maps frequencies of assigned score vectors to one selected alternative. If f is anonymous, then for each report $r \in \mathcal{R}$, every permutation is a valid report as well, i.e., $\sigma(r) \in \mathcal{R}$. We say that a report set $\mathcal{R}$ is *generated by permutations of a single report* if there exists a vector a such that $\mathcal{R} = \{\sigma(a)\}_\sigma$ for all permutations σ on the elements of a; importantly, the report sets for Plurality and Borda fit this characterization. Other general report sets such as partial rankings and vetos, are also amenable to our findings and are usually generated by permutations of a single report.

We often work with two sets of scores $\mathcal{R}$: for Plurality, we let $\mathcal{R} = \mathcal{R}^e := \{e_i : i \in [m]\}$ be the standard basis, which assigns score 1 to a voter's top alternative, and score 0 to every other candidate. When eliciting fully ordinal rankings, we let the report set

[4] Up to scaling, we can intuitively think of $\mathcal{R}$ as being a finite subset of Δ_m.

$\mathcal{R}^\sigma := \{\sigma([m]-\mathbb{1})\}_\sigma$ be all permutations of $\{0,1,\ldots,m-1\}$, assigning score $m-1$ to a candidate's top alternative, $m-2$ to their second favorite, down to 0 score to their least preferred alternative.

Given voter i's preferences $u_i \sim \mu$, we assume voter i chooses to report the score $r \in \mathcal{R}$ minimizing $d(u_i, r)$, where $d(\cdot,\cdot)$ is Euclidean distance. The scoring menu $\mathcal{R}$ induces a function $\Gamma^{\mathcal{R}} : u \mapsto \arg\min_{r\in\mathcal{R}} d(u,r)$ mapping real-valued cardinal utilities to the ordinal preferences submitted by voters with that utility profile. If $\mathcal{R}$ is understood from context, we simply denote this function Γ. [5] Often, we work with the level set $\Gamma_r = \{u \in \Delta_m : r \in \Gamma(u)\}$, which is the set of cardinal preferences leading to a voter submitting $r \in \mathcal{R}$. Each kite in Figure 1 represents a different level set for $\mathcal{R}^e$, and each triangle in Figure 2 represents a different level set for $\mathcal{R}^\sigma$.

Example 2 (Plurality voting). Consider the Plurality voting rule over $m = 3$ alternatives: a, b, and c. A vote for alternative a ascribes scores $(1,0,0)$ to alternatives (a,b,c) respectively; a vote for alternative b ascribes $(0,1,0)$, and a vote for c score $(0,0,1)$. Consider the set of utilities over three outcomes $\Delta_3 = \{u \in \mathbb{R}^3_+ \mid \|u_i\|_1 = 1\}$. The level set $\Gamma_{(1,0,0)} = \{u \in \Delta_3 \mid u_1 \geq u_2 \text{ and } u_1 \geq u_3\}$ is then the set of utility profiles representing voters who most prefer alternative a. For a voter with cardinal utility $(\frac{1}{5}, \frac{1}{2}, \frac{3}{10})$, we have their vote $\Gamma((\frac{1}{5}, \frac{1}{2}, \frac{3}{10})) = \{(0,1,0)\}$, meaning they vote for alternative b. Equivalently, their utility $(\frac{1}{5}, \frac{1}{2}, \frac{3}{10}) \in \Gamma_{(0,1,0)}$ is in the $(0,1,0)$ level set. See Figure 1 for depictions of the level sets for plurality and Figure 2 for Borda and other voting rules that elicit full ordinal rankings over outcomes.

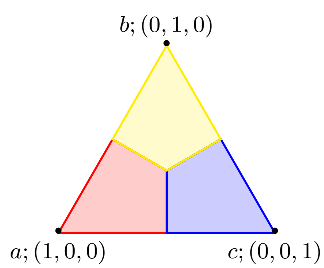

Fig. 1. Plurality: Voters whose utilites lie in the red region (lower left) vote for alternative a by ascribing score $(1,0,0)$, and similar for yellow (top)voters voting for alternative b, and blue (right) voting for alternative c.

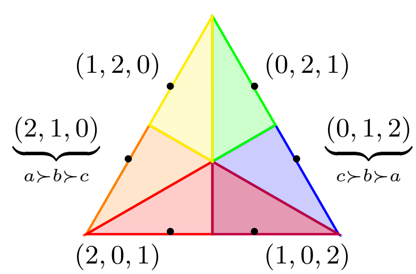

Fig. 2. Fully ordinal preferences: Voters whose utility lies in the orange region vote $a \succ b \succ c$ by ascribing score $(2,1,0)$.

Given a profile of n utilities $\mathbf{u} \in \Delta_m^n$, we consider the induced *histogram* of votes $\#\Gamma^{\mathcal{R}}(\mathbf{u}) \in \mathbb{N}^R$, which are determined by cardinal utility profiles $\mathbf{u}$ and induced by $\Gamma^{\mathcal{R}}$. The r^{th} entry of the histogram represents the number of voters submitting report r. At times, we denote a general histogram of votes by $h \in \mathbb{N}^R$.

[5] We suppose ties are broken uniformly at random between relevant parties, though ties occur with probability 0 if the density μ is sufficiently smooth.

Example 3 (Induced histograms). Consider the Borda voting rule with the report set $\mathcal{R}$ being all $m!$ permutations of $\{0, 1, \ldots, m-1\}$. If $m = 3$ and $n = 15$, a voting profile (histogram) $h = (1, 4, 2, 2, 5, 1) \in \mathbb{N}^{m!}$ implies one voter prefers $a \succ b \succ c$, four voters prefer $a \succ c \succ b$, etc., fixing the order of $\mathcal{R}$. If the voting rule is Plurality, then $h \in \mathbb{N}^m$, and $h = (5, 4, 6)$ means 5 voters have a as their most preferred candidate, 4 have b, and 6 have c as their most preferred.

We generally examine anonymous positional scoring rules $f : \mathbb{N}^R \to [m]$, which are those that select the alternative with highest score given a histogram of score assignments $h \in \mathbb{N}^R$ [26]. We assume ties are broken uniformly at random, as is common in the literature (cf. [11]). Given a report set $\mathcal{R}$ and n voters, let $f_a = \{h \in \mathbb{N}^R \mid a \in f(h), \|h\|_1 = n\}$ denote the set of histograms which lead to alternative a being selected by voting rule f.

Example 4 (Determining the winner of an election). Suppose $m = 3$ and fix n. For alternative a, the set of histograms selecting alternative a according to Plurality voting, $f_a^{plurality} = \{h \in \mathbb{N}^m \mid h_{(1,0,0)} \geq h_{(0,1,0)}$ and $h_{(1,0,0)} \geq h_{(0,0,1)}\}$ is the set of histograms where a receives more votes than b and c. If the voting rule is Veto, then the histograms selecting alternative a $f_a^{veto} = \{h \in \mathbb{N}^m \mid h_{(1,0,1)} \geq h_{(0,1,1)}$ and $h_{(1,1,0)} \geq h_{(0,1,1)}\}$ are those that veto a less than b and c.

Voters have access to a piece of external information $w \in \Delta_m$ and are influenced by it with weight $\alpha \in [0, 1]$. Instead of voting according to their "true" utilities u, voters vote for the report $r \in \mathcal{R}$ minimizing their externality-induced preferences, $d((1-\alpha)u_i + \alpha w, r)$. That is, each voter shifts their preferences α towards w then votes. We are interested in understanding the role of w and α in changing voting outcomes, particularly its effect on social welfare.

Example 5 (Changing vote via w). Suppose a voter has utility $u = (1/2, 9/20, 1/20)$, meaning they very slightly prefer the first outcome the most, then the second, and have almost no utility for the third, i.e., $a \succ bc$. If external information suggests $w = (0, 1/2, 1/2)$ and they are influenced with weight $\alpha = \frac{1}{10}$, then their updated preferences are $(0.45, 0.455, 0.095)$, and they switch their vote from candidate a to candidate b.

Since cardinal preferences are drawn from the unknown density μ, ordinal preferences are typically approximated by measuring the frequency $\frac{|\{i | \Gamma(u_i) = r\}|}{n} \approx \mu(\Gamma_r) = \int_{\Gamma_r} d\mu$. The measure of these level sets serve as a central piece of our analysis, as informational externalities shift the distribution μ in some directed way. Without knowledge of μ, we cannot compute the shifted distribution μ', but we can compute shifts on the coarser measures of level sets (i.e., $\mu(\Gamma_r^{\mathcal{R}})$.) In general, we denote the probability distribution over reports $r \in \mathcal{R}$ with $p := \{\mu(\Gamma_r)\}_{r \in \mathcal{R}}$.

3 A Voting Rule for Eliciting Preferences with Externalities

In practice, algorithm designers do not know the density of voters' utilities μ, and instead approximate the probability a voter submits report r by considering the measure of the report set $\mu(\Gamma_r) = \int_{\Gamma_r} d\mu \approx \frac{|\{i:\Gamma(u_i)=r\}|}{n}$ with only access to n reports $\{\Gamma(u_i)\}_{i\in[n]}$ instead of cardinal utilities $\{u_i\}_{i\in[n]}$. This yields a more practical representation of the density μ, but sacrifices the granular information needed to understand the impacts of the externality w.

In order to understand the effect of informational externalities on voting outcomes, we need to elicit preferences both with and without the presence of these externalities. To this end, we present an intermediary, asymmetric voting rule that enables us to compute true preferences unaffected by externalities. This mechanism modifies the scoring menu $\mathcal{R}$ into a *hypothetical scoring menu* $\mathcal{M}$ so that voting under u with options in $\mathcal{M}$ always yields the same vote as voting under the externality-induced preferences $(1-\alpha)u + \alpha w$ with options in $\mathcal{R}$. With this tool, we are able to elicit the utility profiles u and juxtapose them with the externality-induced preferences with just two queries to voters, instead of eliciting full cardinal utilities.

Theorem 1 provides the scoring menu $\mathcal{M}$, inducing a new voting rule $\Gamma^{\mathcal{M}}$ and bijection $\phi : \mathcal{R} \to \mathcal{M}$ such that $\mathcal{M}$ "mitigates" informational externalities, demonstrated with the equivalence $(1-\alpha)u + \alpha w \in \Gamma_r \iff u \in \Gamma^{\mathcal{M}}_{\phi(r)}$ for all $r \in \mathcal{R}$. That is, for every report $r \in \mathcal{R}$, an externality-induced vote with the standard mechanism yields the report r if and only if a externality-agnostic vote under the perturbed menu also produces vote $\phi(r)$. Through this intermediate voting rule, we elicit slightly more granular information than standard ordinal preferences, but do not require full cardinal preferences. The proof of Theorem 1 is inspired by previous results on property elicitation [21, 22] and power diagrams [4].

Theorem 1. *Given a scoring menu $\mathcal{R}$ inducing $\Gamma^{\mathcal{R}}$, a voter with externality-induced preferences (parameterized by w, α) will vote the same way under $\Gamma^{\mathcal{M}}$ induced by the scoring menu $\mathcal{M} := \{\frac{r-\alpha w}{1-\alpha}\}_r$. That is, $(1-\alpha)u + \alpha w \in \Gamma_r \iff u \in \Gamma^{\mathcal{M}}_{\phi(r)}$ for all $r \in \mathcal{R}$, where $\phi : r \mapsto \frac{r-\alpha w}{1-\alpha}$.*

Proof. Let $v = (1-\alpha)u + \alpha w$. For a score r, we have

$$\begin{aligned} v \in \Gamma_r &\iff d(v,r) \le d(v,r') \qquad \forall r' \neq r \\ &\iff \|r - v\|^2 \le \|r' - v\|^2 \qquad \forall r' \neq r \\ &\iff \left\|\frac{r-\alpha w}{1-\alpha} - u\right\|^2 \le \left\|\frac{r'-\alpha w}{1-\alpha} - u\right\|^2 \qquad \forall r' \neq r \\ &\iff \left\|u - \frac{r-\alpha w}{1-\alpha}\right\|^2 \le \left\|u - \frac{r'-\alpha w}{1-\alpha}\right\|^2 \qquad \forall r' \neq r \\ &\iff u \in \Gamma^{\mathcal{M}}_{\phi(r)} . \end{aligned}$$

Theorem 1 shows that we can use $\mathcal{M}$ as a set of scoring vectors which "mitigates" external information. This allows us to conceptually replace the shift in votes with a shift in the voting rule, and avoid trying to compute a distribution shift for a distribution we do not even know. We benefit as we can now equivalently use the histogram of standard utilities when voting according to $\mathcal{M}$, the histograms $\#\Gamma^{\mathcal{M}}(\mathbf{u}) = \#\Gamma^{\phi(\mathcal{R})}(\mathbf{u})$, which is equal to the histogram of externality-induced preferences according to standard scores $\#\Gamma^{\mathcal{R}}((1-\alpha)\mathbf{u}+\alpha w)$ by Theorem 1. Deriving $\mathcal{M}$ then allows us to empirically estimate the distribution $q := \{\mu(\Gamma_s^{\mathcal{M}})\}_{s\in\mathcal{M}}$ when μ is not known exactly, and does not require eliciting voters' cardinal preferences. In the sequel, we slightly abuse notation and let $\Gamma_r^{\mathcal{M}}$ denote what is technically $\Gamma_{\phi(r)}^{\mathcal{M}}$, which is the set of utilities which would lead voters to submit report r with informational externalities. Observe, since ϕ is defined by α, the mechanism designer is the entity imparting the weight of informational externalities α rather than the voters.[6]

3.1 Externality-Induced Votes Align with w on the Individual Level

We now ask how might an individual's vote change, given an externality $w \in \Delta_m$. Perhaps unsurprisingly, we show that individual votes only move to reports that "better align" with w in this model. We later extend these insights about *individual votes* to *outcomes of voting mechanisms* in § 4. These individual-level results hold for all w and voter preferences u, but we later require mild assumptions on w to characterize changes in population-level voting outcomes.

Proposition 1. *Fix $w \in \Delta_m$ and $\alpha \in [0,1)$, and consider the function $\phi : r \mapsto \frac{r-\alpha w}{1-\alpha}$. Moreover, let $\mathcal{R}$ be a scoring menu generated by one report. For any two reports $s,t \in \mathcal{R}$ such that $d(w,s) \leq d(w,t)$, we have $d(u,s) \leq d(u,t) \implies d(u,\phi(s)) \leq d(u,\phi(t))$.*

Proposition 1 suggests that reports "aligning with w" increase in probability on the individual level: if an voter moves from t to s, then w must "prefer s to t", given by $d(w,s) \leq d(w,t)$. Equivalently, no voter will move from a report s if w prefers the original report over s. This phenomena is demonstrated in Figure 3, where the red cell, corresponding to the vote for alternative a, most preferred by w, is strictly larger under informational externalities. Proposition 1 immediately implies Corollary 1, which states that the set of cardinal utilities yielding the report $r^\star$ that "would be submitted by w" strictly increases.

Corollary 1. *Let an informational externality $w \in \Delta_m$ and weight $\alpha \in (0,1]$ be given and consider the report $\{r^\star\} = \Gamma^{\mathcal{R}}(w)$ for $\mathcal{R}$ generated by one report. Then $\Gamma_{r^\star}^{\mathcal{R}} \subsetneq \Gamma_{r^\star}^{\mathcal{M}}$.*

Under the common *impartial culture* assumption that μ is the uniform density m, this further implies that the probability of any externality-induced preference is ordered according to w. Recall that we do not generally requite the impartial culture assumption, unlike much of the previous literature. As we later frame distributional smoothness relative to the uniform density, this result is helpful later in this study.

Corollary 2. *Let f be an anonymous voting rule. If m is the uniform density, then for all $s,t \in \mathcal{R}$ generated by permutations of a single report, the probability an voter reports s is greater than the probability they report t if and only if w prefers s to t. That is, $m(\Gamma_s^{\mathcal{M}}) \geq m(\Gamma_t^{\mathcal{M}}) \iff d(w,s) \leq d(w,t)$.*

This implies that, when μ is uniform, incorporating external information w leads to the probabilities of alternatives being elected to follow the same order as w if we recursively apply this argument to all pairs of reports s,t.

[6] If α varies for each voter and menus were customized, we can no longer use the assumption that voters are identically distributed. Intuitively, we think of α as an average of individual α_is if α_i is drawn from a single-peaked and roughly symmetric distribution.

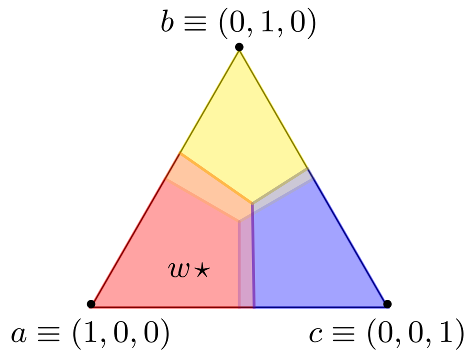

Fig. 3. Plurality votes induced by the externality w yields a menu $\mathcal{M}$ with the level sets of $\Gamma^{\mathcal{M}}$ laid over the level sets of $\Gamma^{\mathcal{R}}$. Utilities in red and yellow intersection (hatched) will vote for candidate b under standard preferences and candidate a with the externality. They are close to indifferent between a and b under standard preferences.

4 Externality-Induced Preferences can Increase the Probability w's Preferred Alternative is Selected

We now provide non-asymptotic bounds on voting in expectation by reducing the problem to a non-asymptotic balls-to-bins problem. This approach yields lower and upper bounds on the probability an alternative a is selected in § 4.1. In this context, we do not compute the exact probability an alternative is selected, but give lower bounds yielding the "worst case" on this probability, and upper bounds giving a "best case" probability an alternative a is selected.

Upon establishing these bounds, we proceed to characterize if and when externality-induced preferences tighten these lower bounds or loosen the given upper bounds in § 4.2. For Plurality and Borda, we show in Corollary 3 that, for the alternative $a^\star$ "most preferred by w," such that $\{a^\star\} = \arg\max_{a'} w_{a'}$, the lower bounds do indeed tighten when the voting rule is Plurality, and give sufficient conditions on w that tighten the lower bounds for Borda. Moreover, if μ is sufficiently smooth, we also show the upper bounds increase in Corollary 4.

4.1 Bounding the Probability Alternative a is Selected via Balls-to-bins

We first examine the probability of different outcomes being reached by a few standard voting rules, though our balls-to-bins approach generalizes. We do this by bounding the probability any given alternative a is selected as the outcome. Observe that, for all (sensible) voting rules f, there is some proportion $c \in (0, 1]$ such that, for all alternatives $a \in [m]$, there exists a set of reports $Q_a \subseteq \mathcal{R}$ such that a c proportion of reports belonging to the set Q_a guarantees that a selected by the voting mechanism. As an example, in Plurality over 3 candidates, the set $Q_a = \{(1, 0, 0)\}$ receiving at least $c = \frac{1}{2}$ of reports implies that outcome a is selected, as demonstrated by the majority criterion. Likewise, for Borda over 3 candidates, we will consider the set $Q_a^\sigma = \{(2, 1, 0), (2, 0, 1)\}$ receiving at least a $c = \frac{m-1}{m}$ fraction of votes. Formally, this condition is written $\sum_{r \in Q_a} h_r > cn \implies f(h) = \{a\}$ is sufficient for a vote to be determined solely by the finite set Q_a of reports, rather than searching over the entire space of votes. If $c = \frac{1}{2}$, this usually reduces to the majority criterion [28, p. 266]. With the report set $\mathcal{R}^\sigma$, we more formally use the set $Q_a^\sigma := \{r \in \mathcal{R}^\sigma : a \text{ is highest ranked alternative}\}$.

With this sufficient condition, we bound the probability alternative a is selected by a reduction to the nonasymptotic balls-to-bins setting[7] studied by Schulte-Geers and Waggoner [31], who give nonasymptotic upper and lower bounds on the probability of the max-loaded bin (of a subset of all bins) having at least k balls. In § 4.2, we show that if voters have externality-induced preferences, these lower bounds tighten for w's most preferred alternative $\{a^\star\} = \arg\max_{a'} w_{a'}$, which demonstrates the influence of the externality w. Throughout this section, we let $p := \{\mu(\Gamma_r)\}_r$ be the probability distribution over coarse reports $\mathcal{R}$.

Plurality First, we use the fact that plurality satisfies the majority criterion as a sufficient condition for an alternative a to be selected by $f^{plurality}$, which we reduce to balls-to-bins. Moreover, by the Mean Value Theorem, we know that some alternative must receive n/m votes in Plurality voting, with which we obtain upper bounds on the probability of any alternative a winning an election. While we are technically concerned with the "sufficient report set" $Q_a = \{e_a\}$ (the basis vector for alternative a), we abbreviate p_{Q_a} to p_a for Plurality.

Proposition 2. *For a density μ defined over $\mathcal{R}^e$, consider $p := \{\mu(\Gamma_r)\}_r$. For any $a \in [m]$, we have*

$$\min\left(\binom{n}{\frac{n}{m}}\|p_a\|_{n/m}^{n/m}, 1 - \sum_{b\neq a}(\Pr[Binom(n, p_b) > n/2])\right) \geq \Pr_{\mathbf{u}}[f^{plurality}(h) = a]$$
$$\geq \Pr[Binom(n, p_a) \geq n/2].$$

Proposition 2 gives both upper and lower bounds on the probability that the Plurality voting rule selects alternative a. Lower bounds are determined by the probability a Binomial sample of n votes returns at least $cn := n/2$ "successes", and the upper bound is given by either the probability some other alternative receives at least $n/2$ votes or the probability that alternative a receives at least $\frac{n}{m}$ votes; a necessary condition for winning the election.

Borda We derive a similar bound for Borda, though this bound is strictly looser than the bound for Plurality, as the number of votes required for the sufficient condition is $c = \frac{m-1}{m}$: strictly greater than the $\frac{n}{2}$ required for Plurality. For large m, this bound loses its meaning as $c \overset{m\to\infty}{\to} 1$, but it is helpful for small m. Part of the reason for this is that we only examine the "sufficient report set" $Q_a^\sigma := \{r \in \mathcal{R}^\sigma : r(a) = m-1\}$ assigning the maximal $m-1$ points to alternative a, neglecting other reports that still view a quite highly, not but the highest.

[7] Throwing n balls into R bins that land in each bin with probability according to p_r, what is the probability the max-loaded bin has $\geq k$ balls inside? e.g., [25]

Proposition 3. *For any $a \in [m]$, density μ, and $c := \frac{(m-1)}{m}$, then*

$$1 - \sum_{b \neq a} \left(\Pr[Binom(n, p_{Q_b^\sigma}) > cn] \right) \geq \Pr_{\mathbf{u}}[f^{borda}(h) = a]$$
$$\geq \Pr_{\mathbf{u}}[Binom(n, \| \sum_{r \in Q_a^\sigma} p_r \|_{cn}) > cn] \, .$$

The proof follows in the same manner as the proof of Proposition 2. Intuitively, Borda rewards alternatives for being highly, but not highest, ranked. This makes it easier for polarizing alternatives "to lose" a Borda election than Plurality. In contrast, Borda makes it easier "to win" by generally appealing to most voters for less polarizing alternatives. This means the bounds are weaker than those given by Plurality since these sufficient conditions are on "winning" and election, rather than "not losing."

Other voting rules Like Plurality, Copeland and Instant Runoff Voting (IRV) also satisfy the majority criterion. Again, there is a set $Q_a \subseteq \mathcal{R}$ such that $\sum_{r \in Q_a} h_r > \frac{n}{2} = cn$, then $f^{Copeland}(h) = f^{IRV}(h) = \{a\}$, and similar bounds can be derived.

Above, our bounds are based on sufficient conditions for an alternative to be declared the winner. However, in Veto voting, we can intuitively think of the winner as being the "least disliked" alternative, which does not lend itself to (nontrivial) sufficient conditions that are independent of other alternatives.

4.2 Tightening Bounds Under Externality-Induced Preferences

While we cannot compute the exact probability that voters select alternative a under voting rule f, § 4.1 gives a range this probability might fall into: the lower bound gives a "worst case" result from the perspective of alternative a, and the upper bound a "best case." The following results state that the "worst case" for the externality's preferred alternative $a^\star$ improves when voters are affected by the externality. Conversely, the "best case" on this probability of selecting $a^\star$ also improves, increasing the upper bound.

Since the bounds in § 4.1 are given by the binomial distribution, we first study the monotonicity of the binomial distribution in the probability of a success. We show in Lemma 1 that if the probability of a success increases, then the probability that a binomial sample has at least k positive instances (meaning $r \in Q_a$) also increases. For Plurality, this suffices to tighten bounds on the probability of selection for the externality's most favored alternative $a^\star = \arg\max_{a'} w(a')$. However, for Borda, we have to restrict the externality w in order for this monotonicity to apply (Lemma 2). While the statements themselves are unsurprising, we are not aware of previous attempts to bound these selection probabilities for generic distributions over ordinal preferences $p := \{\mu(\Gamma_r^{\mathcal{R}})\}_{r \in \mathcal{R}}$ and $q := \{\mu(\Gamma_{\phi(r)}^{\mathcal{M}})\}_{r \in \mathcal{R}}$.

Lemma 1 (Monotonicity of binomial bounds in p). *Suppose $p \leq q$. Then for any $k \in [n]$, we have* $\Pr[Binom(n, p) \geq k] \leq \Pr[Binom(n, q) \geq k]$.

One corollary of Proposition 1 is that when the report set is the set of elementary basis vectors, $\mathcal{R} = \mathcal{R}^e$, then $p_{a^\star} \leq q_{a^\star}$ for all w and α. In turn, Lemma 1 tightens the Plurality lower bounds from Proposition 2 for all $w \in \Delta_m$.

However, when $\mathcal{R} = \mathcal{R}^\sigma$, this is not necessarily the case. For intuition, if w is close to indifferent between two alternatives as its preferred, then there might be some report $s \notin Q_{a^\star}$ that does not rank $a^\star$ as its top choice, yet because of utility over non-top outcomes, we might have $d(w, s) \leq d(w, t)$ for some $t \in Q_{a^\star}$. This opens up the possibility that the probability a voter ascribes the highest score to alternative $a^\star$ decreases, i.e., $\sum_{r \in Q_{a^\star}} p_r > \sum_{r \in Q_{a^\star}} q_r$, as votes might move from t to an "irrelevant report" s. In Lemma 2, we give a sufficient condition on w for the binomial lower bound (the "worst case" for $a^\star$) to increase, illustrated with $m = 3$ alternatives in Figure 4. Intuitively, if w favors one alternative enough, then the reports $r \in \mathcal{R}$ that most align with w are those that assign the highest score to a.

In the sequel, let the function $\min{-k}(\cdot)$ denote the k smallest elements of a given set, where we typically consider $k = (m-1)!$ for the report set $\mathcal{R}^\sigma$.

Lemma 2 (Sufficient condition on w for probability of a report in $Q^\sigma_{a^\star}$ to increase). *Let $\mathcal{R} = \mathcal{R}^\sigma$ and $Q^\sigma_{a^\star} = \min{-k}(\{d(r, w)\}_{r \in \mathcal{R}})$, then $\sum_{r \in Q^\sigma_{a^\star}} q_r \geq \sum_{r \in Q^\sigma_{a^\star}} p_r$.*

Lemma 2 shows that if reports assigning score $m-1$ to alternative $a^\star$ are the k most preferred reports by the externality w, then the probability that any one vote assigns score $m-1$ to alternative a increases. Now we show that this implies a (weak) increase in the probability that alternative $a^\star$ receives a maximal score of $m-1$. For ease of exposition, let $x_{[i]}$ denote the i^{th} largest element of x.

Lemma 3. *Fix report set $\mathcal{R}^\sigma$ and $w \in \Delta_m$, with $a^\star := \arg\max_{a'} w_{a'}$. If $w_{a^\star} \geq (m-1)w_{[2]} + \sum_{i=3}^m w_{[i]}(m - 2(i-i))$, then $Q^\sigma_a := \min{-k}(\{d(w, r)\}_{r \in \mathcal{R}})$ implies $q_{Q^\sigma_{a^\star}} \geq p_{Q^\sigma_{a^\star}}$.*

With $m = 3$, Lemma 3 shows that it suffices for $w_b \leq 1/3$ for all $b \neq a$ in order for the probability that alternative $a^\star$ is the top-ranked to increase. The condition of Lemma 3 is captured by Assumption 4.2. Throughout, we let $w_{[i]}$ denote the i^{th} largest component of the vector w.

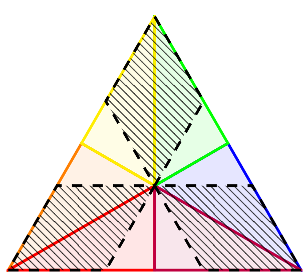

Fig. 4. Hatched regions are subsets of Δ_3 where any w in the hatched region ensures that reports where $a^\star$ is the favorite are precisely the $\min{-2}$ most preferred, satisfying Assumption 4.2. In turn, for any w in the hatched regions, the bounds on the probability that w's most favored alternative is selected increase.

For $m \geq 3$, let w be an externality such that $w_{[1]} \geq w_{[2]}(m-1) + \sum_{i=3}^m w_{[i]}(m - 2i + 2)$.

In essence, we assume the externality w favors one alternative over others strongly enough. If this is true, the probability that a voter ranks that alternative the highest increases. This now enables us to observe that the lower bounds derived in § 4.1 tighten under informational externalities.

Corollary 3. *Let w be an informational externality with $\{a^\star\} = \arg\max_{a'} w_{a'}$. The lower bound given by Proposition 2 tightens for the alternative $a^\star$. Moreover, if $\mathcal{R} = \mathcal{R}^\sigma$ and w satisfies Assumption 4.2, then the lower bound from Proposition 3 tightens.*

Corollary 3 shows that externality-induced preferences can, but do not necessarily, tighten the lower bounds on the probability of a voting rule selects the alternative $a^\star$ most preferred by w. Moreover, the upper bound loosens if μ is sufficiently smooth.

Corollary 4. *Let w be an externality such that $a^\star = \arg\max_{a'} w_{a'}$. Suppose $a^\star$ is the only alternative that increases in probability (i.e., $p_{Q_b} \geq q_{Q_b}$ for all $b \neq a^\star$). Then the upper bounds in Propositions 2 and 3 increase.*

5 Changes in Social Welfare Arising from i Informational Externalities

While Proposition 1 implies that *individuals* will only vote more in line with w, it is well-known that Paretian and non-dictatorial voting rules cannot satisfy independence of irrelevant alternatives [3], and intuitively, this suggests that that externality-induced voting may change votes to upweigh irrelevant alternatives. While an individual vote might increase the score of an irrelevant alternative, it's unclear when and if that will actually change the alternative selected by the voting rule. In order to understand this, we now examine histograms of voting profiles.

For a fixed, finite set of n voters, the set of possible induced histograms is finite, though large. Given a density μ on individual utilities and the induced distribution $p := \{\mu(\Gamma_r)\}$, weight α, and externality w, we examine $\Gamma^{\mathcal{M}}$ and examine the histogram distributions ν and $\nu^{\mathcal{M}}$ such that

$$\nu(f_a) = \sum_{h:a\in f(h)} \frac{1}{|f(h)|} \Pr_{\mathbf{u}}[\#\Gamma^{\mathcal{R}}(\mathbf{u}) = h] = \sum_{h:a\in f(h)} \frac{1}{|f(h)|} \left(\frac{n!}{\prod_r h_r!}\right) \prod_r p_r^{h_r}$$

and likewise for $\nu^{\mathcal{M}}(f_a)$, substituting $\Pr_{\mathbf{u}}[\#\Gamma^{\mathcal{R}}((1-\alpha)\mathbf{u} + \alpha w) = h]$, which is a function of $q := \{\mu(\Gamma^{\mathcal{M}}_{\phi(r)})\}$. The $\frac{1}{|f(h)|}$ term breaks ties uniformly at random.

5.1 Expected Social Welfare Increases when External Information Reflects Social Preferences

We are now equipped to ask how an externality w affects social welfare. We give a sufficient condition for social welfare to increase in expectation in Theorem 2. Then, in Proposition 4 we show that for an externality w roughly in line with the "average utililty" $\mathbb{E}_\mu u$, this sufficient condition is satisfied. In this section, we let $\mathbf{SW}(f, \mathbf{u}) := \sum_i u_i(f(\#\Gamma^{\mathcal{R}}(\mathbf{u})))$ be the total utility attained by the voting outcome when the histogram of reports when preferences are submitted in line with $\mathbf{u}$ (according to $\Gamma^{\mathcal{R}}$, where $\mathcal{R}$ is induced by f). Throughout, we let $\Delta\mathbf{SW}(f, \mathbf{u})$ denote the change in welfare arising from informational externalities $\Delta\mathbf{SW}(f, \mathbf{u}) := \mathbf{SW}(f, (1-\alpha)\mathbf{u} + \alpha w) - \mathbf{SW}(f, \mathbf{u})$, and we ask when the welfare increases in expectation, i.e., $\mathbb{E}[\Delta\mathbf{SW}(f, \mathbf{u}) \mid w, \alpha] \geq 0$. That is, when does social welfare increase (in expectation) under informational externalities?

Theorem 2. *Let w, α, μ be given. Denote the set* $\mathbf{inc} := \{a \in [m] : \nu^{\mathcal{M}}(f_a) - \nu(f_a) \geq 0\}$ *be the set of alternatives that weakly increase in probability of being selected under informational externalities, and* $\mathbf{dec} := [m] \setminus \mathbf{inc}$ *be the set of alternatives that decrease. If* $\max_{a \in \mathbf{dec}} (\mathbb{E}\mu)_a \leq \min_{a \in \mathbf{inc}} (\mathbb{E}\mu)_a$*, then* $\mathbb{E}_{\mathbf{u}}[\Delta \mathbf{SW}(f, \mathbf{u}) \mid w, \alpha] \geq 0$.

Proof. For ease of exposition, let $v := \mathbb{E}_{u \sim \mu} u$. First, observe that since social welfare is additive and preferences are drawn iid, we can restrict our focus to v instead of $\mathbf{u}$.

$$
\begin{aligned}
\mathbb{E}_{\mathbf{u} \sim \mu^n}[\Delta \mathbf{SW}(f, \mathbf{u}) \mid w, \alpha] &= \mathbb{E}_{\mathbf{u} \sim \mu^n}[\sum_{i \in [n]} \sum_{a \in [m]} u_i(a) \left(\Pr_{\mathbf{u}}[f(h(\mathbf{u}, \Gamma^{\mathcal{M}})) = a \mid w] - \Pr_{\mathbf{u}}[f((h(\mathbf{u}, \Gamma)) = a] \right) \\
&= \mathbb{E}_{\mathbf{u} \sim \mu^n}[\sum_{i \in [n]} \langle u_i, \nu^{\mathcal{M}}(f) - \nu(f) \rangle] \\
&= n \mathbb{E}_{u \sim \mu}[\langle u, \nu^{\mathcal{M}}(f) - \nu(f) \rangle] \\
&= n \langle v, \nu^{\mathcal{M}}(f) - \nu(f) \rangle \geq 0 \iff \langle v, \nu^{\mathcal{M}}(f) - \nu(f) \rangle \geq 0
\end{aligned}
$$

Now we argue that under the assumptions, that $\langle v, \nu^{\mathcal{M}}(f) - \nu(f) \rangle \geq 0$.

First, observe that since ν and $\nu^{\mathcal{M}}$ are probability distributions (both subject to affine constraints), $\sum_a \nu^{\mathcal{M}}(f_a) - \sum_a \nu(f_a) = 0$, and therefore

$$
\sum_{a \in \mathbf{inc}} (\nu^{\mathcal{M}}(f_a) - \nu(f_a)) = - \sum_{a \in \mathbf{dec}} (\nu^{\mathcal{M}}(f_a) - \nu(f_a)) \quad = \sum_{a \in \mathbf{dec}} (\nu(f_a) - \nu^{\mathcal{M}}(f_a)). \tag{1}
$$

This yields

$$
\langle v, \nu^{\mathcal{M}}(f) - \nu(f) \rangle \geq 0 \iff \sum_{a \in \mathbf{inc}} v_a(\nu^{\mathcal{M}}(f_a) - \nu(f_a)) \geq \sum_{a \in \mathbf{dec}} v_a(\nu(f_a) - \nu^{\mathcal{M}}(f_a))
$$

We can show the latter inequality, as

$$
\begin{aligned}
\sum_{a \in \mathbf{inc}} v_a(\nu^{\mathcal{M}}(f_a) - \nu(f_a)) &\geq \min_{a \in \mathbf{inc}} v_a \sum_{a \in \mathbf{inc}} (\nu^{\mathcal{M}}(f_a) - \nu(f_a)) \\
&= \min_{a \in \mathbf{inc}} v_a \sum_{a \in \mathbf{dec}} (\nu(f_a) - \nu^{\mathcal{M}}(f_a)) \qquad \text{Eq. (1)} \\
&\geq \max_{a \in \mathbf{dec}} v_a \sum_{a \in \mathbf{dec}} (\nu(f_a) - \nu^{\mathcal{M}}(f_a)) \qquad \text{by assumption} \\
&\geq \sum_{a \in \mathbf{dec}} v_a(\nu(f_a) - \nu^{\mathcal{M}}(f_a))
\end{aligned}
$$

Intuitively, Theorem 2 shows that expected social welfare increases if the externality w is at least a good approximation of the true expected utility $\mathbb{E}\mu$. However, it is unclear how reasonable the assumptions of Theorem 2 are. To better understand these assumptions, we turn to the notion of ϵ-uniformity (Definition 1), and bound how large ϵ can be while still yielding the assumptions of Theorem 2.

In particular, if the density of preferences μ has a small total variation distance (TV) from the uniform density, then the assumptions of Theorem 2 are satisfied, yielding Proposition 4.

Definition 1 **(ϵ-uniform density).** *A density μ is ϵ-uniform if,* $\sup_{A\in\mathcal{B}} \left|\int_A d\mu - \int_A dm\right| \leq \epsilon$*. That is, $TV(\mu, m) \leq \epsilon$, where m is the uniform density and TV denotes the total variation distance.*

Our notion of ϵ-uniformness is a much stricter condition than necessary, but its generality allows us to draw conclusions about *any* w that is ordered roughly according to the "average utility" $\mathbb{E}_\mu u$, and is largely agnostic to the choice of voting rule.

We now give a series of results which together help us show that, for a small enough ϵ on the ϵ-uniformness, informational externalities crudely aligned with social welfare lead to expected increases in social welfare. While the conditions on ϵ seem tight, this is in part because of the generality of our results: Proposition 4 states that social welfare increases in expectation as long as (a) f is an anonymous positional scoring rule and (b) w and $\mathbb{E}\mu$ have the same ordering over alternatives: a strictly weaker condition than being equal.

Proposition 4. *Suppose f is an anonymous positional scoring rule, and w and $\mathbb{E}\mu$ have the same order. Moreover, suppose μ is ϵ-uniform for $\epsilon \leq \min_{s,t\in\mathcal{R}} \frac{|m(\Gamma_s^{\mathcal{M}})-m(\Gamma_t^{\mathcal{M}})|}{2}$ and $\frac{\sum_{k=1}^{n}\binom{n}{k}(R\epsilon)^k}{m} \leq \min_{a,b} \frac{|\nu(f_a)-\nu(f_b)|}{2}$, then expected social welfare increases.*

While we show that expected social welfare increases if w roughly aligns with social welfare, an adversarial externality w that is ordered in the reverse of $\mathbb{E}\mu$ will generally lead to a decrease in social welfare. This demonstrates the influence exhibited by the externality w.

5.2 Bounding the Probability that Externalities Decrease Social Welfare

Upon learning that welfare increases in expectation with an externality roughly in line with social welfare, a natural next question is to understand the probability the social welfare *decreases*, even when expected to increase. Since many voting mechanisms are often run in one-shot settings, understanding the probability that social welfare decreases is important. To this end, we now give an upper bound on the probability social welfare deceases which is agnostic to the choice of w. Informally, Theorem 3 upper bounds the probability that social welfare decreases under informational externalities by the expected exponent of the change in social welfare.

Theorem 3. *Given a density μ such that preferences are drawn i.i.d. from μ, the externality w and weight $\alpha \in [0,1]$, then* $\Pr_{\mathbf{u}}[\Delta\mathbf{SW}(f,\mathbf{u}) < 0] \leq \mathbb{E}[\exp(\mathbf{SW}(f,\mathbf{u}) - \mathbf{SW}(f,(1-\alpha)\mathbf{u}+\alpha w)]$.

Proof. For notational simplicity, fix any $\epsilon \in (0, \min_{a,b:w_a>w_b}(w_a - w_b))$ and let $V := -\Delta\mathbf{SW}(f,\mathbf{u}) = \mathbf{SW}(f,(1-\alpha)\mathbf{u}+\alpha w) - \mathbf{SW}(f,\mathbf{u})$.

$$
\begin{aligned}
\Pr[\Delta\mathbf{SW}(f,\mathbf{u}) < 0] &= \Pr[-\Delta\mathbf{SW}(f,\mathbf{u}) \geq \epsilon] && \text{Finite alternatives } m \\
&\leq M_V(t)e^{-t\epsilon} \qquad \forall t > 0 && \text{Chernoff bound} \\
&= \mathbb{E}_V[\exp(tV)]\exp(-t\epsilon) \qquad \forall t > 0 && \text{definition of MGF} \\
&\leq \mathbb{E}_V[\exp(tV)] && \forall t > 0, -t\epsilon < 0 \implies \exp(-t\epsilon) \in (0,1) \\
&\leq \mathbb{E}_V[\exp(V)] && \text{choose } t = 1 \\
&= \mathbb{E}_{\mathbf{u}}[\exp(\mathbf{SW}(f,\mathbf{u}) - \mathbf{SW}(f,(1-\alpha)u+\alpha w)]
\end{aligned}
$$

Observe this bound is only meaningful (i.e., it is tighter than 1) if externality-induced voting increases social welfare in expectation.

6 Discussion

Future Work While our work so far has considered only one informational externality w, the polarized nature of media content practically leads us to wonder if our results naturally extend to a vector of informational externalities $\vec{w}$ through some clever aggregation of $\vec{w}$ to one "summarizing point" $\bar{w}$. In short: the answer is no, at least for $m \geq 3$ alternatives.

The problem that arises with the attempt to apply our results "out of the box" is the timing of aggregation. In essence, trying to apply our results requires aggregating a vector $\vec{w}$ into one $\bar{w}$ to analyze the relative position of the aggregated $\bar{w}$. However, in reality, aggregation of different externalities happens at the histogram level (once preferences have already been elicited). One can concatenate the histograms induced by multiple externalities $\vec{w}$ by evaluating the level sets of histograms and further understanding the distributional shifts on ν, rather than distributional shifts on μ, as covered by most of our results. While we remain optimistic for the mechanism introduced in § 3 to provide insight into histogram-level aggregations on ν, we leave these extensions as open work.

Conclusion This work provides a methodology to examine the robustness of voting mechanisms to external information. In applying our methodolgy for eliciting externality-induced preferences, we focus on the case where the externality w is roughly aligned with social welfare, unsurprisingly finding that imposing an externality aligned with social welfare tends to increase social welfare in expectation. Moreover, we use a balls-to-bins approach to give upper and lower bounds on the probability that any one alternative is selected by the original voting rule, and show that, for the alternative a most preferred by the externality, this lower bound tightens and the upper bound can loosen under informational externalities.

Acknowledgements We would like to thank Bailey Flanigan for helpful comments in scoping out and situating this work, and Niclas Böhmer for helpful edits and feedback along the way. This material is based upon work supported by the National Science Foundation under Award No. 2202898 and by National Science Foundation and Amazon under Award No. 2147187.

Competing Interests The authors have no competing interests to declare that are relevant to the content of this article.

Bibliography

[1] Elliot Anshelevich, Onkar Bhardwaj, Edith Elkind, John Postl, and Piotr Skowron. Approximating optimal social choice under metric preferences. *Artificial Intelligence*, 264:27–51, 2018.

[2] Jose Apesteguia, Miguel A Ballester, and Rosa Ferrer. On the justice of decision rules. *The Review of Economic Studies*, 78(1):1–16, 2011.

[3] Kenneth J Arrow. *Social choice and individual values*, volume 12. Yale university press, 1951.

[4] F. Aurenhammer. Power Diagrams: Properties, Algorithms and Applications. *SIAM Journal on Computing*, 16(1):78–96, February 1987.

[5] Amirreza Bagheridelouee. Metric Distortion Under Public-Spirited Voting. *New Zealand*, 2024.

[6] Dorothea Baumeister and Tobias Hogrebe. Complexity of Election Evaluation and Probabilistic Robustness. *New Zealand*, 2020.

[7] Mark Bedaywi, Bailey Flanigan, Mohamad Latifian, and Nisarg Shah. The distortion of public-spirited participatory budgeting. 2023.

[8] Niclas Boehmer, Robert Bredereck, Piotr Faliszewski, and Rolf Niedermeier. On the Robustness of Winners: Counting Briberies in Elections, October 2020. arXiv:2010.09678 [cs].

[9] Niclas Boehmer, Robert Bredereck, Piotr Faliszewski, and Rolf Niedermeier. A Quantitative and Qualitative Analysis of the Robustness of (Real-World) Election Winners. In *Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization*, EAAMO '22, pages 1–10, New York, NY, USA, October 2022. Association for Computing Machinery.

[10] Niclas Boehmer, Piotr Faliszewski, Łukasz Janeczko, and Andrzej Kaczmarczyk. Robustness of Participatory Budgeting Outcomes: Complexity and Experiments, May 2023. arXiv:2305.08125 [cs].

[11] Craig Boutilier, Ioannis Caragiannis, Simi Haber, Tyler Lu, Ariel D. Procaccia, and Or Sheffet. Optimal social choice functions: A utilitarian view. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, EC '12, page 197–214, New York, NY, USA, 2012. Association for Computing Machinery.

[12] Canadian Press. France's election: 'strategic' voting among montreal's french citizens, 2022.

[13] Ioannis Caragiannis, Swaprava Nath, Ariel D Procaccia, and Nisarg Shah. Subset selection via implicit utilitarian voting. *Journal of Artificial Intelligence Research*, 58:123–152, 2017.

[14] Moses Charikar, Prasanna Ramakrishnan, Kangning Wang, and Hongxun Wu. Breaking the metric voting distortion barrier. *arXiv preprint arXiv:2306.17838*, 2023.

[15] Po-An Chen, Bart De Keijzer, David Kempe, and Guido Schäfer. Altruism and its impact on the price of anarchy. *ACM Transactions on Economics and Computation (TEAC)*, 2(4):1–45, 2014.

[16] Edith Elkind, Piotr Faliszewski, and Arkadii Slinko. Swap Bribery. In Marios Mavronicolas and Vicky G. Papadopoulou, editors, *Algorithmic Game Theory*, pages 299–310, Berlin, Heidelberg, 2009. Springer.

[17] Bailey Flanigan, Ariel D Procaccia, and Sven Wang. Distortion under public-spirited voting. *ACM Conference on Economics and Computation*, 2023.

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

[19] Glory M. Givi, Robin Delabays, Matthieu Jacquemet, and Philippe Jacquod. On the robustness of democratic electoral processes to computational propaganda. *Scientific Reports*, 14(1):193, January 2024. Publisher: Nature Publishing Group.

[20] Libby Jenke and Scott A. Huettel. Voter Preferences Reflect a Competition Between Policy and Identity. *Frontiers in Psychology*, 11:566020, October 2020.

[21] Nicolas Lambert and Yoav Shoham. Eliciting truthful answers to multiple-choice questions. In *Proceedings of the 10th ACM conference on Electronic commerce*, pages 109–118, 2009.

[22] Nicolas S Lambert. Elicitation and evaluation of statistical forecasts. 2018.

[23] John Ledyard. Public goods: A survey of experimental research, 1997.

[24] Aki Lehtinen and Jaakko Kuorikoski. Unrealistic assumptions in rational choice theory. *Philosophy of the Social Sciences*, 37(2):115–138, 2007.

[25] Michael Mitzenmacher, Andréa Richa, and Ramesh Sitaraman. The power of two random choices: A survey of techniques and results. 2001.

[26] Elchanan Mossel, Ariel D Procaccia, and Miklós Z Rácz. A smooth transition from powerlessness to absolute power. *Journal of Artificial Intelligence Research*, 48:923–951, 2013.

[27] Citizens' Assemblies – Democracy that works.

[28] James Roland Pennock and John W Chapman. *Due Process: Nomos XVIII*, volume 30. NYU Press, 1977.

[29] Shawn W. Rosenberg, Lisa Bohan, Patrick McCafferty, and Kevin Harris. The Image and the Vote: The Effect of Candidate Presentation on Voter Preference. *American Journal of Political Science*, 30(1):108–127, 1986. Publisher: [Midwest Political Science Association, Wiley].

[30] 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.

[31] Ernst Schulte-Geers and Bo Waggoner. Balls and bins – simple concentration bounds, 2022.

[32] Dmitry Shiryaev, Lan Yu, and Edith Elkind. On Elections with Robust Winners. *AAMAS*, 2013.

[33] Piotr Skowron and Edith Elkind. Social choice under metric preferences: Scoring rules and stv. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, 2017.

[34] Ilia Tsetlin, Michel Regenwetter, and Bernard Grofman. The impartial culture maximizes the probability of majority cycles. *Social Choice and Welfare*, 21:387–398, 2003.

[35] E. Glen Weyl. The robustness of quadratic voting. *Public Choice*, 172(1):75–107, July 2017.