Expressive Voting: Modeling a Voter's Decision to Vote

Dominik Klein

1 Introduction

The "paradox of voting" was identified by Anthony Downs in his seminal book *An Economic Theory of Democracy* [4]. In a large election, there is almost no chance that an individual vote will have any effect on the outcome of the election. Thus, faced with the decision of whether or not to vote, if there is any cost associated with the act of voting, then the only *rational* choice[1] for a voter is not to participate. Much of the voting theory literature bypasses this problem by assuming that there is no cost to voting or that "strategic" considerations do not play a role in the voters' decision to vote. See [5] for a discussion of the different approaches to this problem.

However there is a second approach, prominently presented by Brennan and Lomasky, [2] whose line of argumentation roughly runs like this: A priori, utility could be associated with the outcome on an action or the performing of the act itsself. For modelling many cases of human choice behaviour these two choices do not make any observable behavioral difference (e.g. in ordering wine A rather than wine B vs. drinking wine A rather than wine B). Consequently, revealed preference analysis has developed a general preference for outcome based utilites. In voting situations however there turns out to be a crucial difference between the two kinds of utility attachments - precisely because casting a certain ballot does *not* guarantee that one gets the desired outcome, nor is it any probable that the outcome changes at all.

Obviously, attaching utilities to sole voting acts avoids the "paradox of voting", since one is guaranteed to receive the output utility associated with the voting statement made. To introduce some terminology: We call the voting behaviour induced by utilites on the outcome *instrumental voting*, since it takes the ballot as an instrument to bring about a certain political state. On the other hand, the voting behaviour induced by utility attachments on acts is called *expressive voting*, since it reflects the voter's original preferences over parties. Brennan and Lomasky [2] argue that expressive reasoning about voting actually occurs and plays a major role in voting behaviour. Arguably, almost any realistic reasoning about whom to vote involves a superposition of the two principles, where the relative weights between these two depend upon certain aspects of the situation[2]. Thus, in evaluating different voting procedures, one needs to analyze on how these compare from an instrumental as well as an expressive standpoint.

While the instrumental approach of voting is widely studied in decision theory, the import of expressive voting in comparing voting systems is not very well understood. (see [2] for some analysis). In this paper, we explore an interesting new approach to this problem found in a recent paper by Enriqueta Aragones, Itzhak Gilboa and Andrew Weiss [1], henceforth called the AGW-approach. The basic idea of their model is that there are certain factual topics on the agenda and the agents' utilites for the individual parties are derived by their positions on these topics. The structure of the paper is the following: In paragraph 2 we present the general model introduced by Aragnoes, Gilboa and Weiss and their semantics of

[1] In the sense that taking into account the cost of the act of voting, the expected utility of voting for a preferred candidate (or set of candidates) will be negative.

[2] See [2] pp. 40-46

approval voting. Paragraph 3 then contains a critique of their semantics of approval voting, followed by our approach to this problem together with a defense of the choices we made in paragraph 4. Our results are in paragraph 5, follwed by some extensions in paragraph 6. Finally, the appendix contains all proofs and calculations.

2 The model

In this section, we briefly describe the basic model of the AGW-approach. Suppose that $T = \{1, \ldots, m\}$ is a set of parties, or candidates. Each party $j \in T$ is characterized by its positions on the various issues of concern $I = \{1, \ldots, n\}$. To this end each party $j \in T$ is associated with a vector $\mathbf{p}^j \in [-1; 1]^n$ giving j's positions on each of the issues. Then, $p_i^j \in [-1; 1]$ is the degree to which candidate[3] j supports issue i. For notational convenience, we will use $\mathbf{p}$ with decorations to denote parties or candidates, where $\mathbf{v}$ with all its variants denotes voters. A voter's ballot is represented by a vector $\mathbf{x} = (x_1, \ldots, x_m) \in \mathbb{R}_+^m$. Thus, *abstention* corresponds to the zero vector $\mathbf{0}$ (in which $\mathbf{0}_j = 0$ for each j). A voting systems consists of a set $F \subset \mathbb{R}_+^m$ of *feasible* values for $\mathbf{x}$ together with an aggregation rule for these feasible values[4]

The key idea of the AGW-model is that each ballot $\mathbf{x}$ is associated with a "statement" giving a position on each issue. The statement made by a ballot $\mathbf{x} \in \mathbb{R}_+^m$ on a topic i is $\sum_{j \in T} x_j v_i^j \in \mathbb{R}$. Thus the entire statement is the vector

$$\left(\sum_{j \in T} x_j v_i^j \right)_{i \leq n} \in \mathbb{R}^n$$

As for parties, also each voter is represented by a vector $\mathbf{v} \in [-1; 1]^n$ representing his position on the various topics. The decision problem faced by voter k is then to find the ballot that makes a statement as close as possible to her actual position. That is to find the ballot $\mathbf{x}$ that minimizes the Euclidean distance from the statement made by $\mathbf{x}$ to the voter's ideal position w. More precisely, if F is the set of feasible ballots, then voter $\mathbf{v}$ must solve the following minimization problem:

$$\min_{\mathbf{x} \in F} \mathrm{dist}(\mathbf{v}, \sum_{j \in T} x_j \mathbf{p}^j)$$

where $\mathrm{dist}(x, y) = \sqrt{\sum (x_i - y_i)^2}$ is the usual Euclidean distance. If the solution to this minimization problem is $x = \mathbf{0} \in \mathbb{R}^m$, then the voter will abstain. Thus, a voter's choice to abstain is due to an inability to express herself in the voting system rather than any cost associated with voting. The main contribution of AGW [1] is a rigorous comparison of majority rule and approval voting using the above model of voting. The definition of the two voting systems runs as follows:

Majority rule: voters select a single candidate. In this case, the degree of support is $x^j = 1$ for the selected party and $x^j = 0$ for the others. Thus, the feasible ballots are $F^M = \{0\} \cup \{e^j\}_{j \leq m}$ where e^j is the vector with 1 in the jth position.

---

[3] In the AGW approach it is always assumed that $\mathbf{v}_i^j \in \{-1; 1\}$, i.e. parties have extreme positions. Briefly, their argument is that political discourse moves parties to extreme stances

[4] For instance in borda count the feasible $\mathbf{x}$ are permutations of $\{1 \ldots m\}$, while in majority voting the feasible $\mathbf{x}$ have at most one entry with value 1 and all other entries 0.

Approval voting: voters select any subset I of candidates and the voter with the most approvals wins the election. Given the above assumptions, the approval ballots are:

$$F^A = \left\{ x^I \in \mathbb{R}^m \mid I \subseteq \{1, \dots, m\} \text{ and } x^I = \frac{1}{|I|} \sum_{j \in I} e^j \right\}$$

where $x^\emptyset$ is the zero vector. Thus, approving of a set I corresponds to supporting each candidate in I to degree $\frac{1}{|I|}$. Given the semantics, this amounts to averaging opinions in I for each topic.

Obviously there are more statements available to the voter under approval voting than majority rule (i.e., $F^M \subseteq F^A$). Expanding the set of feasible statements can only increase participation in elections. Building on this intuition, AGW prove two results that illustrate the ways that a richer set of statements can lead to increased participation in elections.

3 Criticism of the Model

We introduce two minimal claims about the way a voter reasons about his expressive vote. We then show that Gilboa's semantics violates both of these claims. The first claim is that the way a voter reasons about a voting system is closely connected to the way the system actually works. That is public discourse about an election and its outcome shapes the way a voter reasons about his expressive statements. An electoral system that produces single winners does not involve concepts of a coalition or a compromise, and so doesn't public discourse about this system. Thus it is, and this is the first criticism, highly unlikely to assume that a typical voter start reasoning in these terms.

For the second criticism recall that in expressive voting situations, a voter's payoff does not depend upon the actual outcome, but only on his vote. Nevertheless, we hold that potential outcomes do have a certain influence on an expressive vote. We hold the following to be a reasonable criterion to be fulfilled by an account of expressive voting: a voter $\mathbf{v}$ approves of the outcome of any election in which every voter submits the same ballot as $\mathbf{v}$ himself, or equivalently a voter approves of the outcome of any election where he is the single voter.

We show that the AGW semantics also violates this condition. For the intuition behind the counterexample assume that there is a set of moderate parties and to opposing extremist parties. It might so happen that the position of a moderate voter $\mathbf{v}$ is exactly the average between two extremist parties - even though every moderate party is closer to him than each of the extremists. Given the AGW semantics of approval voting, $\mathbf{v}$ would have to approve of exactly the two extremist parties. Now if every voter submitted the same ballot as $\mathbf{v}$ all votes would go to the two extreminst parties, and thus one of the two (depending upon some tie-breaking rule) would get into office. But by our assumptions this is the outcome $\mathbf{v}$ dislikes most among all possible outcomes.

Since the AGW approach models positions on individual topics rather than degrees of extremism, we cannot directly translate the above story into a formal counterexample. However the following values mimick the main features of the above example:

Example: An election is based upon some issues $t_1 \dots t_9$. Voter $\mathbf{v}$ is interested in the first four of them with weights $\frac{1}{3}, -\frac{1}{3}, \frac{1}{3}, -\frac{1}{3}$ respectively. All other topics receive weight 0. The two extremist parties are are $\mathbf{e}_+$ assigning 1 to every topic and $\mathbf{e}_-$ assigning -1 to every topic. Every other party $\mathbf{p}_i$ assign weights $1, -1, 1, -1$ to the first four topics and 1 to all remaining topics. Then the setup is as claimed above, i.e. all moderate parties $\mathbf{p}_i$ are closer to $\mathbf{v}$ than $\mathbf{e}_+$ and $\mathbf{e}_-$, but $\{\mathbf{e}_+, \mathbf{e}-\}$ is the approval set chosen by $\mathbf{v}$. See the appendix for details.

As a third criticism we claim that the idealizations used in the AGW model are only reasonable in a single-winner election: We present voters by vectors anywhere in $[-1;1]^n$ exactly to allow them a differential weighting of alternatives. For a party in a single winner system AGW restricts itsself to $\mathbf{v} \in \{-1;1\}^n$ for two reasons. The first has been mentioned above: that political discourse can move parties to an extrem stand on all topics[5]. But this is only conlclusive in conjunction with a second hypothesis: That the single winner will implement its positions on both topics simply because it is the only party in charge. Thus there is no point in comparing topics w.r.t. to relevance, since the party stands for implementing all of the topics it stands for. In contrast, in thinking about a coalitional semantics we need relative weights to predict the position of coalitions. Observations of coalitional agreements indicate that the position of a coalition is not decided by taking straight averages (i.e. majority vote) over the individual topics. Rather, every coalitional has some core topics that it is unwilling to make any compromise at. Other topics outside these core areas are just all too open for reconsiderations. Thus the AGW approach seems to be conflating intuitions coming from a single winner system and intuitions coming from a coalitional system.

While we are sympathetic with the general approach of modelling parties as positions on topics, the above criticism gives good reasons to be hesitant in accepting the semantics given by AGW. In the following paragraph we give a competing semantics based on the same general model. We will show that our semantics avoids all of the above mentioned criticism while validating similar results as the AGW approach

4 Our model

For majority voting we take the same semantics as AGW. While majority is about choosing the best possible alternative (if it is sufficiently attractive), approval voting is about deciding on how much one is willig to deviate from the optimum and then going with every party that deviates less than this amount. We assume that for every topic a party will either implement a policy or its converse, thus each $\mathbf{p}_i \in \{-1;1\}$ for each topic i. The payoffs for a voter $\mathbf{v}$ on topic i, given that party $\mathbf{p}$ comes to power is thus either $|v_i|$ or $-|v_i|$, depending on whether the signs of v_i and p_i agree or not. Formally the payoff is given by

$$\begin{aligned} |v_i| &\text{ iff } v_i \cdot p_i \geq 0 \\ -|v_i| &\text{ iff } v_i \cdot p_i < 0 \end{aligned}$$

Combining the two the payoff of $\mathbf{v}$ on topic i is given by $p_i \cdot v_i$. Thus, we can formally define our semantics for approval voting:

Approval voting: Fix some $k \in [-1;1]$, the *approval coefficient.* A voter $\mathbf{v}$ approves of all parties $\mathbf{p}$ that satisfy:

$$\sum p_i v_i \geq k \cdot \sum |v_i| \tag{1}$$

Using that $\sum p_i v_i$ is the standard scalar product $\mathbf{p} \cdot \mathbf{v}$ and that $\sum |v_i|$ is the 1-norm $|\mathbf{v}|_1$, equation 1 translates to

$$\mathbf{p} \cdot \mathbf{v} \geq k|\mathbf{v}|_1 \tag{2}$$

Some remarks to this definition:

[5]Of course, different topics might be *stressed* differently in campaigning. The only assumption AGW makes is that parties have positions on all topics. See [3] for a competing model

- Typically $k \geq 0$, that is an the party and the agent agree on more topics (weighted) than they disagree
- An alternative interpretation for k is in terms of weighted percentual agreement: Agent $\mathbf{v}$ and party $\mathbf{p}$ agree on at least $\frac{1+k}{2}$ of the topics weighted by the appropriate $\mathbf{v}_i$.
- The above definition has an interesting geometric interpretation: For a vector $x \in \mathbb{R}^n$ and some nonnegative real number α let $\mathcal{C}(x, \alpha)$ be the cone of all vectors in $\mathbb{R}^n - \{0\}$ such that the angle between x and y is at most α. Then we have the following lemma

  Lemma 4.1. *Let $\mathbf{v}$ be a voter and let k be as in the definition of approval voting. Then there is some angle α depending upon n, k and $\mathbf{v}$ such that for each party $\mathbf{p}$ holds*

  $$\mathbf{p} \in C(\mathbf{v}, \alpha) \Leftrightarrow \mathbf{p} \cdot \mathbf{v} \geq k|\mathbf{v}|_1$$

  Furthermore, α satisfies $\arccos(k) \leq \alpha \leq \arccos(\frac{k}{\sqrt{n}})$

  Thus if we interpret the vector $\mathbf{v}$ as giving a voter's general direction, the above formula says that a voter approves of all parties that lie in roughly the same direction.

Before giving our mathematical results, we note that our semantics avoids all of the above mentioned criticism: In the semantics for approval voting presented here the voter reasons about parties individually. That is, he evaluates a party by the *what-if* state that would obtain if a party came to power. Consequentally, the proposed way of reasoning is sufficiently similar to the way the voting system works. This adresses the first criticism. For the second criticism note that a voter only votes for parties he approves of individually. Thus any election in which every voter casts the same ballot as $\mathbf{v}$ will produce an outcome $\mathbf{v}$ approves of individually.

5 Results

The first question to answer is: Some agent might mistakenly use the toolbox for approval voting in majority situations. It is an important question whether the two semantics given (majority vote vs. approval vote restricted to one-party ballots) give the same result. In the Gilboa case this is fairly straightforward. The following lemma shows that it also holds true for our case. We stipulate that a single-vote voter in a approval ballot votes for the party $\mathbf{p}$ that maximizes the quotient $\frac{\sum v_i \cdot p_i}{\sum |v_i|}$. Or to say it in other words: The voter $\mathbf{v}$ goes with the party $\mathbf{p}$ that minimizes the angle between $\mathbf{v}$ and $\mathbf{p}$.

Lemma 5.1. *Let $\mathbf{v}$ be a voter, let P be a set of parties and let $\mathbf{p}^* \in P$ be a party. Then we have:*

$$\text{dist}(\mathbf{p}^*, \mathbf{v}) = \min_{\mathbf{p} \in P} \text{dist}(\mathbf{p}, \mathbf{v}) \Leftrightarrow \frac{\sum v_i p_i^*}{\sum |p_i^*|} = \max_{p \in P} \frac{\sum v_i p_i}{\sum |p_i|}$$

Next, we show that our approach satisfies the same properties as the AGW approach. Bascially, we examine how many parties are needed to offer everyone an alternative he prefers to abstaining. Not very suprisingly, the result depend heavily upon the approval coefficient k. For $k = 0$, i.e. $\alpha = 90°$ our results are very similar to those obtained by AGW. The following two theorems roughly correspond to theorems 1 and 2 of AGW. Theorem 5.2 is much more general than the corresponding theorem 1 of AGW, since we do not want to assume that voters value all topics equally. Since for majority voting our semantics is the same as AGW, their results also hold for our approach. The following analysis is entirely about approval voting.

The first question we ask is: How many strategically distributed parties are needed to ensure that every voter finds something he prefers to abstaining in an approval system

Theorem 5.2. *i) If $k \leq 0$ two parties are enough to ensure that every voter approves of at least one party.*
ii) If $k > 0$ the number of parties needed to ensure that no (possible) voter abstains is exponential in the number of topics.

Now we are interested in the a priori chance that a voter finds some party he approves of, given that there are n parties. To this end we construct a random party by throwing a fair coin to determine its position on each topic T_i seperately and independently from each other. We are interested in the probability that given and an approval coefficient k and n such random parties (where n also is the number of topics) some voter $\mathbf{v}$ finds at least one partie he approves. This probability depends upon the vector $\mathbf{v}$ so let the $P_A(n,n,k)$ denote[6] the minimum over all $\mathbf{v}$ of the probability that $\mathbf{v}$ approves of at least one out of n random parties with approval coefficient k.

Theorem 5.3. *For $k \leq 0$ we have $\lim_{n\to\infty} P_A(n,n,k) = 1$. On the other hand for $k \in (0;1]$ the converse holds: $\lim_{n\to\infty} P_A(n,n,k) = 0$*

Remark: The pessimistic result $\lim_{n\to\infty} P_A(n,n,k) = 0$ depends crucially upon the choice of $\mathbf{v}$. For many choices of $\mathbf{v}$, for instance $\mathbf{v} \approx \pm e_i$, where e_i is the i-th unit vector in $\mathbb{R}^n$, we have $\lim_{n\to\infty} P_A(n,n,k) = 1$ for all $k \in [-1;1]$. Thus a universally interested voter is harder to accidentaly satisfy than somebody who is only intereted in very few topics. Given that the latter might hold true for many voters, our result seems to be an overly pessimistic worst case result.

Remark: Alternatively, we could have defined approval voting by replacing the $\geq$ by a $>$ in equation 1. This would only slightly influence our results: 5.2 would still hold with 2 replaced by $2n$. For theorem 5.3 It is easy to see that this would not influence the result, since for any non-zero voter $\mathbf{v}$ the set $\{\mathbf{x} \in [-1;1]^n] | \mathbf{b} \cdot \mathbf{x} = 0\}$ has at most cardinality 2^{n-1}.

6 Extensions

We propose two further applications of our framework. The first is the extension to a more general class of voting systems, namely range voting as an extension of approval voting. The second is about incorporating focus and focus dynamics into the framework.

i) Range voting. Range voting refers to a family of voting system. The underlying idea behind all of these is that voters are asked to grade candidates with a given linear scale of grades. The systems then differ in how to turn the ballots into a result[7], but this question is irrelevant for our approach since we assume the voter's payoff to be completely determined by his sole vote. Of course, approval voting is a special case of range voting, where the admissible grades are *approve* and *disapprove*. We will show that a good deal of our analysis caries over to range voting. Assume that the grades used are $g_1 \ldots g_n$ with g_1 being the worst and g_n the best. Instead of the one approval coefficient from the approval voting there

[6]Notation is chosen in line with the original paper. To be precise: $P_n(n,m,k)$ denotes the minimum of all agents $\mathbf{v}$of the probability that in an m-topic election with n parties $\mathbf{v}$ approves of at least one party with approval coefficient $\mathbf{k}$

[7]For instance if the grade scale is numerical using means vs. using averages

is a set of grade requirements $-1 = t_1 \leq \ldots \leq t_n \in [-1; 1]$. The grade of voter $\mathbf{v}$ for some party $\mathbf{p}$ is then given by

$$grade(\mathbf{v}, \mathbf{p}) := \max \{i | \mathbf{v} \cdot \mathbf{p} \geq t_i |\mathbf{v}|_1\}$$

Translated into the language of cones, this means that a voter is associated with a sequence $\mathcal{C}(\mathbf{v}, \alpha_1) \supseteq \ldots \supseteq \mathcal{C}(\mathbf{v}, \alpha_n)$ of narrower and narrower cones, where the first is the entire space. The grade of a party is then given by the highest index i such that $\mathbf{p} \in \mathcal{C}(\mathbf{v}, \alpha_i)$ or equivalently by the number of cones $\mathbf{p}$ is in. It seems worthy to remark that we do not assume the t_i to be the same for various voters. All we assume is that each voter has a set of grade requirements that he uses to assign the grades to parties - and of course the set of grades available to voters is the same for everyone. Obviously, not every voter has to make use of the entire spectrum of grades available. For instance a voter restricting himself to only use the two extreme grades g_1 and g_n is described by the additional constraint $t_2 = \ldots = t_n$. On the other hand, approval voting can be seen as a special case of range voting with the grade requiremtnes $t_{disapproval} = -1, t_{approval} = k$.

As above, we are interested in how many parties are needed to give an arbitrary voter the incentive to cast his vote. We use a slightly stronger criterion than in the approval case: We ask how many parties are needed to give voter $\mathbf{v}$ a relevant difference he wants to express in grades, that is some parties $\mathbf{p}$ and $\mathbf{p}'$ receiving different grades. Interestingly, this depends crucially upon whether the voters take one of their grade requirements to express the fact *I agree with* $\mathbf{p}$ *more often than I disagree*, i.e. $t_i = 0$. The following are counterparts of theorems 5.2 and 5.3 above. Again, the first is about strategically positioning parties on the map.

Theorem 6.1. *Assume that every voter has some i with $t_i = 0$. Then $2n$ parties are enough to ensure that every (possible) voter finds two parties he grades differently.*

For the next theorem, we have to adapt the definition of $P(n, n, k)$ to the new setting. Denote the set of grade requirements $t_1 \ldots t_n$ by $\mathbf{t}$ and let $P_A^{\mathbf{t}}(n, n, k)$ denote the minimum over all $\mathbf{v}$ using grade requirements $\mathbf{t}$ of the probability that $\mathbf{v}$ finds two parties that he assigns different grades to.

Theorem 6.2. *If there is some i with $t_i = 0$ then $\lim_{n\to\infty} P_A^{\mathbf{t}}(n, n, k) = 1$. On the other hand if there is no such i then $\lim_{n\to\infty} P_A^{\mathbf{t}}(n, n, k) = 0$*

ii) Public discourse. Until now, our basic model presents voters' preferences as static. Of course, good arguments might convince voters to switch their stands over time, though this process is as slow as rare in occurence. There is a second type of dynamics much more relevant for predicting electoral outcomes. Will party A be able to convince voters to see the election as a vote on economic policy? Can party B ensure that voters think about gun control when making their electoral decision? It is as important for a party to have competitive opinions on individual topics, as it is important to move these topics into public *focus*. In the following, we examine how effects of focussing can be incorporated into our framework. First, we observe that for a voter $\mathbf{v}$, some party $\mathbf{p}$ and a fixed set of topics, the degree of $\mathbf{v}$'s approval to $\mathbf{p}$ does not change if some topics that $\mathbf{v}$ does not care about are removed from the agenda. This is the content of the following lemma.

Lemma 6.3. *For approval voting the following holds: Let $K \subseteq \{1 \ldots n\}$ be a subset of the topics. For any party $\mathbf{p}$ let $\mathbf{p}_K$ be the restriction of $\mathbf{p}$ to K and similarly for voters. Fix some $\mathbf{v}$ and assume that $\mathbf{v}_i = 0$ for all $i \notin K$. Then*

$$\frac{\mathbf{v} \cdot \mathbf{p}}{|\mathbf{v}|_1} = \frac{\mathbf{v}_K \cdot \mathbf{p}_K}{|\mathbf{v}_K|_1}$$

Arguably, a change in focus does not change a voter's general attitude, that is the sign, of some particular position. It does however change the length $|v_i|$ of the entries as a measure for relative importance of topics. For this end, we define a focus-changing matrix as:

Definition 6.4. *A focus matrix is a diagonal matrix*[8] $A \in [0;1]^{n\times n}$. *Voter* $\mathbf{v}$*'s position after focus change with* A *is denoted by* $\mathbf{v}_A := A \circ \mathbf{v}$.

For further discussion denote the diagonal entries of A by $a_1 \dots a_n$, thus $\mathbf{v}_A = (a_i v_i)_{i\leq n}$ Note that the restriction to some subset $K \subseteq \{1, \dots, n\}$ is exactly given by the matrix A_K with $a_i = 1$ for $i \in K$ and all other entries zero, i.e. the projection to K. That is in the terminology of the above lemma: $\mathbf{v}_K = \mathbf{v}_{A_K}$.

The genesis of public attention is a complex matter. Arguably focus is influenced by the macroscopic situation, external events and many more.But it is also shaped by news coverage, the content of electoral campaigns and other events that are at least partially under the control of parties or symphasizing groups. The question arises of which utterances a party *should* make in order to maximize their electoral chances. This has for instance be dealt with by Parikh and Dean in [3], though their model is not about focus, but about the question on how much information a party should reveal about its intended policy. We suggest a different view: Party campaining is primarily not about dispensing information about one's plans, but about bringing one's core strengh into focus. Thus, we model the intended effect of campaining by playing a certain focus matrix $A_{\mathbf{p}}$. As mentioned above, several focuses will be in the arena: The focus matrices played by the various parties, together with some focuses emerging from the general situation, external events and many more. These can then be combined into a general focus by taking the straight average of all focus matrices. A general question to be elaborated in further work is. Which focus matrix should a party play, i.e. on which topics should it campaign to maximize its appproval. For now note that since $\mathbf{v}_A$ satisfies our definition of a voter again, theorems 5.2 and 6.1 give conditions on how many parties are needed to ensure that no voter abstains *after every possible focus change.*

7 Discussion

We embrace the idea of interpreting voting as expressions of opinions. This reinterpretation of voting as an expressive language comes along with the need for a semantics for this language. We identify certain drawbacks of the AGW approach and offer a different semantics avoiding these drawbacks. We show that depending open the approval coefficient, our semantics allows to show similar results as in the original paper. We further show that our approach can be extended to range voting. We also hint at ways at how focus and focus dynamics can be incorporated into the framework.

References

[1] Enriqueta Aragones, Itzhak Gilboa, and Andrew Weiss. Making statements and approval voting. *Voting Theory and Decission*, 71:461–472, 2011.

[2] Geoffrey Brennan and Loren Lomasky. *Democracy and Decision – The Pure Theory of Electoral Choice.* Cambridge University Press, 1993.

[3] Walter Dean and Rohit Parikh. The logic of campaigning. In *Logic and Its Applications*, pages 38–49. Springer, 2011.

[8] i.e. all entries outside the diagonal are zero

[4] Anthony Downs. *An Economic Theory of Democracy.* Harper and Row, 1957.

[5] Timothy J. Feddersen. Rational choice theory and the paradox of not voting. *Voting Theory and Decission*, 18:99–112, 2004.

8 Proofs

We start by showing that example 3 satisfies all properties claimes. In particular we have to show that *i)* $\operatorname{dist}(\mathbf{v}, \mathbf{p}) < \operatorname{dist}(\mathbf{v}, \mathbf{e}_\pm)$ and *ii)* that $\{\mathbf{e}_+, \mathbf{e}_-\}$ is the coalition approved by $\mathbf{v}$.

For *i)* observe that

$$\operatorname{dist}(\mathbf{v}, \mathbf{p}) = \sqrt{4 \cdot \left(\frac{2}{3}\right)^2 + 5} = \sqrt{\frac{16}{9} + 5} \text{ and}$$
$$\operatorname{dist}(\mathbf{v}, \mathbf{e}_*) = \sqrt{2 \cdot \left(\frac{2}{3}\right)^1 + 2 \cdot \left(\frac{4}{3}\right)^2 + 5} = \sqrt{\frac{24}{9} + 5} \text{ for } * \in \{+, -\}$$

thus $\mathbf{e}_\pm$ are indeed the most extremist parties.

For *ii)* observe that

$$\operatorname{dist}(\frac{1}{2}(\mathbf{e}_+ + \mathbf{e}_-), v) = \frac{2}{3}$$

To see that this is the closest coalition we first show that any coalition C containing three or more members has a distance of at least $\frac{\sqrt{5}}{3}$ from $\mathbf{v}$. For any such coalition the last five entries of C are all at least $\frac{1}{3}$ (with the minimum reached if C consists of exactly three entries, one of them being $\mathbf{e}_-$). Thus $\operatorname{dist}(C, \mathbf{v}) \geq \frac{\sqrt{5}}{3}$. A similar argument shows that for $C' = \{\mathbf{e}_+, \mathbf{p}\}$ holds $\operatorname{dist}(C', \mathbf{v}) \geq \sqrt{5}$. Finally, for the coalition $\mathbf{C}''\{\mathbf{e}_-, \mathbf{p}\}$ we have

$$\operatorname{dist}(C', \mathbf{v}) = \sqrt{2 \cdot \left(\frac{1}{3}\right)^2 + 2 \cdot \left(\frac{2}{3}\right)} = \frac{\sqrt{10}}{3}$$

thus finishing the proof.

Proof of lemma 4.1. For $x, y \in \mathbb{R}^n$ the angle α between x and y is described by the following well-known equation

$$\frac{x \cdot y}{|x|_2 |y|_2} = \cos \alpha \tag{3}$$

where $|\mathbf{x}|_2 = \sqrt{\sum x_i^2}$ denotes the euclidean length. On the other hand inequality 2 can be transformed to

$$\frac{\mathbf{v} \cdot \mathbf{p}}{|\mathbf{v}|_1} \geq k$$
$$\Leftrightarrow \frac{\mathbf{v} \cdot \mathbf{p}}{|\mathbf{v}|_2 \sqrt{n}} \geq \frac{k}{\sqrt{n}} \frac{|\mathbf{v}_1|}{|\mathbf{v}_2|}$$

Since $|\mathbf{p}|_2 = \sqrt{\sum_i 1} = \sqrt{n}$. This is exactly equation 3 for

$$\alpha = \arccos(\frac{k}{\sqrt{n}} \frac{|\mathbf{v}_1|}{|\mathbf{v}_2|}).$$

The last claim follows from the inequality

$$|x|_2 \leq |x|_1 \leq \sqrt{n}|x|_2$$

for all $x \in \mathbb{R}^n$. $\square$

Proof of lemma 5.1. Recall that $p_i \in \{-1; 1\}$ for each topic $i \in N$. Fix a voter $\mathbf{v}$. For any party $\mathbf{p}$ let $U_{\mathbf{p}} \subseteq \{1 \dots N\}$ be defined by:

$$i \in U_{\mathbf{p}} \Leftrightarrow v_i \cdot p_i < 0$$

Thus $U_{\mathbf{p}}$ is the set of indices where the sign of $\mathbf{v}$ and $\mathbf{p}$ disagrees. Now we have:

$$\begin{aligned}
\operatorname{dist}(\mathbf{v}, \mathbf{p}) &= \sqrt{\sum_i (v_i - p_i)^2} \\
&= \sqrt{n + \sum_i v_i^2 - 2\sum_i v_i p_i} \\
&= \sqrt{n + \sum_i v_i^2 - 2\sum_i |v_i| + 4\sum_{i \in U_{\mathbf{p}}} |v_i|}
\end{aligned}$$

Observe that only the last term depends on $\mathbf{p}$. Thus we have for any $\mathbf{p}, \mathbf{p}' \in P$:

$$\operatorname{dist}(\mathbf{p}, \mathbf{v}) \leq \operatorname{dist}(\mathbf{p}', \mathbf{v}) \Leftrightarrow \sum_{i \in U_{\mathbf{p}}} |v_i| \leq \sum_{i \in U_{\mathbf{p}'}} |v_i|$$

On the other hand we have:

$$\sum_i v_i p_i = \sum_i |v_i| - 2\sum_{i \in U_{\mathbf{p}}} |v_i|$$

Thus also:

$$\frac{\sum v_i p_i}{\sum |p_i|} \geq \frac{\sum v_i p_i'}{\sum |p_i'|} \Leftrightarrow \sum_{i \in U_{\mathbf{p}}} |v_i| \leq \sum_{i \in U_{\mathbf{p}}'} |v_i|$$

$\square$

Before we can prove theorems 5.2 and 5.3 we need the following lemma:

Lemma 8.1. *Let $m \in \mathbb{N}$. Then we for any natural number n:*

$$\frac{\sum_{k=\lceil n(\frac{1}{2}+\frac{1}{2m})\rceil}^{n} \binom{n}{k}}{2^n} \leq 2\left(\left(1 + \frac{1}{2m}\right)^{-1}\right)^n \tag{4}$$

Proof. For notational convenience we assume n to be even. First we show that for any natural number $i \in [0, \frac{n}{2m}]$ we have that

$$\binom{n}{\frac{n}{2} + i} \geq \left(1 + \frac{1}{2m}\right)^n \binom{n}{\frac{n}{2} + \lceil \frac{n}{2m} \rceil + i} \tag{5}$$

To this end observe that

$$
\begin{aligned}
&\frac{\binom{n}{\frac{n}{2}+i}}{\binom{n}{\frac{n}{2}+\lceil\frac{n}{2m}\rceil+i}}\\
=&\frac{(\frac{n}{2}+\lceil\frac{n}{2m}\rceil+i)!(\frac{n}{2}-\lceil\frac{n}{2m}\rceil-i)!}{(\frac{n}{2}-i)!(\frac{n}{2}+i)!}\\
=&\frac{(\frac{n}{2}+i+1)\cdot(\frac{n}{2}+i+2)\cdot\ldots\cdot(\frac{n}{2}+\lceil\frac{n}{2m}\rceil+i)}{(\frac{n}{2}-\lceil\frac{n}{2m}\rceil-i+1)\cdot(\frac{n}{2}-\lceil\frac{n}{2m}\rceil-i+2)\cdot\ldots\cdot(\frac{n}{2}-i)}\\
=&\frac{\frac{n}{2}+1+i}{(\frac{n}{2}-\lceil\frac{n}{2m}\rceil-i)}\cdot\ldots\cdot\frac{\frac{n}{2}+\lceil\frac{n}{2m}\rceil+i}{\frac{n}{2}-i}
\end{aligned}
$$

Now it is easy to see that each of the quotients in the last formula is larger than $1+\frac{1}{2m}$, thus the entire product is larger than $(1+\frac{1}{2m})^n$ and 5 holds. In the following, let $\alpha := \left((1+\frac{1}{2m})^{-1}\right)^n$

Iteratedly applying 5 gives us for all natural numbers j with $0 \leq j < \lceil\frac{n}{2m}\rceil$:

$$
\sum_{i=1}^{m}\binom{n}{\frac{n}{2}+j+i\lceil\frac{n}{2m}\rceil} \leq \sum_{i=1}^{m}\alpha^i\binom{n}{\frac{n}{2}+j} \leq \frac{\alpha}{1-\alpha}\binom{n}{\frac{n}{2}+j} \leq 2\alpha\binom{n}{\frac{n}{2}+j}
$$

where the last inequality holds since $\alpha < \frac{1}{2}$. In particular we have

$$
\begin{aligned}
&\sum_{k=\lceil n(\frac{1}{2}+\frac{1}{2m})\rceil}^{n}\binom{n}{k} = \sum_{i=1}^{m}\sum_{j=0}^{\lceil\frac{n}{2m}\rceil-1}\binom{n}{\frac{n}{2}+j+i\lceil\frac{n}{2m}\rceil}\\
\leq&2\alpha\sum_{j=0}^{\lceil\frac{n}{2m}\rceil-1}\binom{n}{\frac{n}{2}+j} < 2\alpha\sum_{j=0}^{n}\binom{n}{j}
\end{aligned}
$$

Resubstituting $\alpha = (1+\frac{1}{2m})^{-n}$ gives us

$$
\frac{\sum_{k=\lceil n(\frac{1}{2}+\frac{1}{2m})\rceil}^{n}\binom{n}{k}}{2^n} \leq 2\left(1+\frac{1}{2m}\right)^{-n}
$$

□

Proof of theorem 5.2. For i) observe that $\mathbf{p}_1 := (1,1,\ldots 1)$ and $\mathbf{p}_2 := -\mathbf{p}_1$ have the property that for any voter $\mathbf{v}$ at least one of the two statements $\mathbf{p}_1\cdot\mathbf{v} \geq 0$ and $\mathbf{p}_2\cdot\mathbf{v} \geq 0$ holds. Thus each voter approves of at least one of these two parties.
ii). Let $\mathbb{V} := \{-1;1\}^n$ be the set of voters whose have extreme positions on every single topic. We will show that the number of parties needed to ensure that every member of $\mathbb{V}$ votes is exponential in n. Fix some natural number $\frac{1}{m} \leq k$. Since the number of parties some voter $\mathbf{v}$ approves of is decreasing in k it suffices to show the theorem with $k = \frac{1}{m}$. Observe that for any party $\mathbf{p}$ and any voter $\mathbf{v} \in \mathbb{V}$ holds:

$$
\mathbf{v}\cdot\mathbf{p} \geq \frac{1}{m}|\mathbf{v}|_1 \Leftrightarrow |\{i|v_i = p_i\}| \geq \frac{n}{2}+\frac{n}{2m}
$$

Since for any party $\mathbf{p}$ and any $l \in \mathbb{N}$

$$
|\,\{\mathbf{v} \in \mathbb{V} : |\{i : v_i = p_i\}| = l\}\,| = \binom{n}{l}
$$

this implies that each party $\mathbf{p}$ can be be approved of by at most $\sum_{k=\lceil n(\frac{1}{2}+\frac{1}{2m})\rceil}^{n} \binom{n}{k}$ many members of $\mathbb{V}$. Since $|\mathbb{V}| = 2^n$ this implies that the number of parties needed to make sure that no member of $\mathbb{V}$ abstains is at least

$$\frac{2^n}{\sum_{k=\lceil n(\frac{1}{2}+\frac{1}{2m})\rceil}^{n} \binom{n}{k}}$$

By lemma 8.1 this quotient is at least as large as $\frac{1}{2}\left(1+\frac{1}{2m}\right)^n$ in particular it is exponential in n. Since 2^n parties are enough to ensure that everybody votes the number of parties needed cannot be worse than exponential. $\square$

Proof of theorem 5.3: Fix a voter $\mathbf{v}$ Observe that for $k = 0$ and any party $\mathbf{p}$ at least one of the following two holds: $\mathbf{v} \cdot \mathbf{p} \leq 0 \cdot |\mathbf{v}|_1$ or $\mathbf{v} \cdot \mathbf{p} \geq 0 \cdot |\mathbf{v}|_1$. Let $\mathcal{P} = \{-1; 1\}^n$ be the set of all possible parties

$$\frac{|\{\mathbf{p} \in \mathcal{P} | \mathbf{p} \cdot \mathbf{v} \geq 0\}|}{|\mathcal{K}|} \geq \frac{1}{2}$$

Since picking a random party is the same as randomly drawing a party from $\mathcal{P}$, the chance that a random party $\mathbf{p}$ satisfies $\mathbf{p} \cdot \mathbf{x} \geq 0$ is at least one half. Thus the chance that $\mathbf{v}$ approves of none of n random parties is at most $\frac{1}{2}^n$, thus $P(n, n, 0) \to 1$. Obviously, this implies $P(n, n, k) \to 1$ for any $k \leq 0$
Since $P(n, n, k)$ is monotonous in k, it suffices to show that $P(n, n, \frac{1}{m}) \to 0$ for any natural number m. let $\mathbf{v} = (1, 1, \ldots)$ be a voter who fully approves of all topics and let $m \in \mathbb{N}$. Observe that for any party $\mathbf{p}$ holds:

$$\mathbf{v} \cdot \mathbf{p} \geq \frac{1}{m}|\mathbf{v}|_1 \Leftrightarrow |\{i | p_i = 1\}| \geq \frac{n}{2} + \frac{n}{2m}$$

Thus for the uniform distribution $\mathbb{P}$ over $\mathcal{P}$ we have

$$\mathbb{P}(\mathbf{v} \cdot \mathbf{p} \geq \frac{1}{m}|\mathbf{v}|_1) = \frac{\sum_{k=\lceil n(\frac{1}{2}+\frac{1}{2m})\rceil}^{n} \binom{n}{k}}{2^n}$$

As above lemma 8.1 yields that

$$\mathbb{P}(\mathbf{v} \cdot \mathbf{p} \geq \frac{1}{m}|\mathbf{v}|_1) \leq 2\left(\left(1+\frac{1}{2m}\right)^{-1}\right)^n$$

Thus $P(n, n, \frac{1}{m}) \leq 1 - (1 - 2\left(1+\frac{1}{2m}\right)^n)^n$. It is a general fact $(1 - kx^n)^n \to 1$ for any $x \in (0, 1)$ and $k \in \mathbb{R}$, thus $P(n, n, \frac{1}{m}) \to 0$ as claimed.

$\square$

Proof of theorem 6.1. Fix a voter $\mathbf{v}$ and let i such that $t_i = 0$. Then by 5.2 and the remark following it $2n$ parties are enough to guarantee that there is at least one party getting grade larger or equal to i. Applying the remark to the voter $\mathbf{v}$ shows that there is also some party getting grade at most $i - 1$. $\square$

Proof of theorem 6.2. First assume that there is some i such that $t_i = 0$. Then applying theorem 5.3 the probability that at least one out of n random parties gets grade at least i goes to 1. Applying 5.3 to $-\mathbf{v}$ also the probability that a party gets grade at most $i - 1$ goes to 1. In particular, the probability for there being two parties receiving different grade assignments goes to 1, this proves the first part. For the second part assume that there is no such i. Let i_0 be such that for all $t_i < 0$ for all $i \leq i_0$ and $t_i > 0$ for all $i > i_0$. Then applying 5.3 with $k = t_{i_0}$ (if defined) yields that the probability that no party gets grade larger than i_0 goes towards 0. Applying 5.3 to $-\mathbf{v}$ yields that also the probability for parties getting grade lower than i_0 goes to zero, thus the probability of all parties getting the same grade i_0 goes towards 1. $\square$

Proof of lemma 6.3. Observe that in 1 all summands indexed by some i with $v_i = 0$ vanish on both sides of the equation. Thus

$$\frac{\sum_{i \leq n} p_i v_i}{\sum_{i \leq n} |v_i|} = \frac{\sum_{i \in K} p_i v_i}{\sum_{i \in K} |v_i|}$$

$\square$