

---

# Vote Manipulation in the Presence of Multiple Sincere Ballots

---

Ulle Endriss

Institute for Logic, Language and Computation  
University of Amsterdam

## Abstract

A classical result in voting theory, the Gibbard-Satterthwaite Theorem, states that for any non-dictatorial voting rule for choosing between three or more candidates, there will be situations that give voters an incentive to *manipulate* by not reporting their true preferences. However, this theorem does not immediately apply to all voting rules that are used in practice. For instance, it makes the implicit assumption that there is a *unique* way of casting a sincere vote, for any given preference ordering over candidates. Approval voting is an important voting rule that does not satisfy this condition. In approval voting, a ballot consists of the names of any subset of the set of candidates standing; these are the candidates the voter *approves* of. The candidate receiving the most approvals wins. A ballot is considered *sincere* if the voter prefers any of the approved candidates over any of the disapproved candidates. In this paper, we explore to what extent the presence of multiple sincere ballots allows us to circumvent the Gibbard-Satterthwaite Theorem. Our results show that there are several interesting settings in which no voter will have an incentive not to vote by means of *some* sincere ballot.

## 1 INTRODUCTION

A central issue in voting theory concerns the question whether a voting rule has the property that it will never give a voter an incentive to cast a vote that does not sincerely reflect his true preferences over candidates. For instance, in the context of the commonly used *plurality rule* (Taylor, 2005), a voter may well have an incentive to vote for his second favourite can-

didate rather than his first, namely if he knows or believes (maybe after having studied polling information) that the first candidate has no chance of winning anyway, while the second is likely to tie with, say, his least favourite candidate. An ideal voting rule would not require voters to strategise in this manner. The seminal result in the field, the Gibbard-Satterthwaite Theorem (Gibbard, 1973; Satterthwaite, 1975), states that, for *any* non-dictatorial voting rule for choosing between three or more candidates, there will be situations that give voters an incentive to manipulate by not reporting their true preferences. However, in the form it is usually stated (e.g. Taylor, 2005), the Gibbard-Satterthwaite Theorem does not (at least not immediately) apply to a number of voting rules used in practice. There are two reasons for this:

- Firstly, the theorem assumes that voting amounts to reporting a *full* preference ordering over all candidates. The plurality rule, for instance, does not satisfy this condition, as it requires voters to submit ballots with just a single name on them.<sup>1</sup>
- Secondly, the statement of the theorem presupposes that there is a *unique* way of casting a sincere vote, for any given preference ordering over candidates. *Approval voting* (Brams and Fishburn, 1978) is an important voting rule that does not satisfy this condition. In approval voting, a ballot consists of the names of any subset of the set of candidates standing; these are the candidates the voter *approves* of. The candidate receiving the most approvals wins. A ballot is considered *sincere* if the voter prefers any of the approved candidates over any of the disapproved candidates. Hence, there will be multiple sincere ballots for any given preference ordering.

---

<sup>1</sup>This form of balloting limits the ways in which a voter can strategise, which conceivably could affect the validity of the theorem. It applies nevertheless, but strictly speaking that needs to be checked separately.

The aim of this paper is to understand how the presence of multiple sincere ballots affects the voters' incentives to manipulate. We do this within the framework of approval voting.

As observed already by Niemi, a feature of approval voting that may be considered problematic is that “*the existence of multiple sincere strategies almost begs voters to behave strategically*” (Niemi, 1984). That is, even if a voter *wants* to be honest, approval voting does not prescribe a clear course of action. In this paper, we shall contrast this potentially critical aspect of having multiple sincere ballots with a rather positive one. We show that also voters who are not honest *per se* will often not have any incentive to vote insincerely, due to the availability of an alternative sincere ballot yielding just as good (or even better) an outcome.

The remainder of the paper is organised as follows. Section 2 briefly defines approval voting and introduces some of the notation we require. If several candidates receive the same (maximal) number of approvals, we have to decide on a suitable rule for breaking ties. This tie-breaking rule, and possibly also other factors, will determine how a voter rates a *preliminary* election outcome (a set of top candidates, before ties are broken). To be able to assess whether a voter has an incentive to manipulate, we need to extend voter preferences over individual candidates to preferences over sets of candidates. Section 3 gives a set of five axioms for such preference orderings over sets of candidates, which we take to reflect the minimum requirements any such ordering has to satisfy. Our results on manipulation are presented in Section 4. We show that if we make either certain (additional) assumptions on voter preferences over sets of candidates, or if we restrict attention to scenarios with certain numbers of candidates (or both), then we are able to prove that approval voting is *not* prone to manipulation in the sense that no voter will ever have an incentive not to vote by means of *some* sincere ballot. We conclude with a brief discussion of the bigger picture and an outlook on possible future work.

## 2 PRELIMINARIES

In this section, we briefly introduce the system of approval voting (Brams and Fishburn, 1978), and explain some of the notation used in this paper.

### 2.1 Approval Voting

Suppose a finite number of *voters* have to elect a *single winner* from a pool of  $m$  *candidates*. In approval voting, a *ballot* is a subset of (the names of) the set of candidates. To vote, each voter submits such a ballot.

A voter is said to *approve* of the candidates included in his ballot, and to *disapprove* of all the remaining ones.<sup>2</sup> The winner of the election is determined by summing up the approvals for each candidate. If there is a unique candidate with a maximal number of approvals, that candidate wins. If there are several, the winner is selected from amongst those with maximal approval using a suitable *tie-breaking rule*. The subject of tie-breaking will be discussed in Section 3.

Each voter is assumed to have a *preference ordering*  $\preceq$  over the set of candidates. We take each  $\preceq$  to be anti-symmetric, transitive, and total.<sup>3</sup> A ballot by a given voter is called *sincere* if all the approved candidates are ranked above all the disapproved candidates according to that voter's preference ordering  $\preceq$ . Otherwise, the ballot is called *insincere*. The underlying intuition is that only such ballots “without gaps” (and including the voter's top choice) are accurate reflections of the voter's true preferences.<sup>4</sup>

### 2.2 Notation for Candidate Sets and Ballots

We use a simplified notation to describe *sets of candidates*, from the point of view of a particular voter. For instance, if there are 4 candidates, then 421 denotes the set including the most preferred candidate (4) as well as the two least preferred candidates (21). To describe *ballots*, we use the same kind of notation,

<sup>2</sup>Submitting either the full set of candidates or the empty set as a ballot amounts to *abstaining*; neither will have any effect on the outcome of an election. To simplify presentation, we do not allow for the empty ballot.

<sup>3</sup>Antisymmetry, in particular, means that there can be no ties between individual candidates (according to the preferences of individual voters).

<sup>4</sup>Our definition of sincere ballots in approval voting is standard and intuitive. But for voting rules *in general*, the question of how to define the notion of *sincere ballot* is interesting and far from trivial (Parikh and Pacuit, 2005). The simple route to take would be to adopt a *normative* approach: define a function mapping any actual preference ordering to the correct (set of) sincere ballot(s). Often, there will be overwhelming cause for a particular choice of function. For instance, if the voting rule requires reporting a *full* preference ordering, then the most obvious choice is to ask that the only sincere ballot for  $\preceq$  be  $\preceq$  itself. An alternative approach would be to define a sincere ballot as the ballot that will give the best expected outcome for the voter, in the absence of any knowledge of how the others are going to vote. Of course, turning this basic idea into a workable model will not be easy (we would most likely require *some* sort of probability distribution over possible tallies, and it seems difficult not to get such “world knowledge” mixed up with knowledge regarding the other voters' actions). Parikh and Pacuit (2005) discuss these issues at length and, in the face of such difficulties of finding a reasonable general definition of sincere ballot, propose and study the alternative notion of *safe ballot*. A safe ballot is a ballot that guarantees the voter an outcome that is never worse than abstaining (and strictly better in some cases).

but enclosed in square brackets. For instance, in the case of 5 candidates, casting the ballot [51] means approving of one’s most and least preferred candidate (an insincere ballot). The full set of sincere ballots for 3 candidates, for instance, would be [3], [32], [321].

### 3 PREFERENCES OVER SETS OF CANDIDATES

Suppose all voters have cast their ballots. We call a candidate a *pre-winner* if she is one of the candidates receiving the most approvals. If there is more than one pre-winner the actual winner will be picked from the set of pre-winners using a *tie-breaking rule*. A natural choice for such a rule would be to draw a winner from amongst the pre-winners using a uniform probability distribution. However, at this stage we do not want to commit to any specific tie-breaking rule.

Naturally, a voter’s preferences over sets of pre-winners are likely to depend, at least in part, on the tie-breaking rule applied. For instance, a voter may (internally) assign a utility of  $m-1$  to the candidate he ranks the highest,  $m-2$  to his second favourite, and so on. If the tie-breaking rule is known to the voter, he will be able to compute the expected utility for each set and base his preferences over pre-winners on that measure. However, again, at this stage we are not going to make any such limiting assumptions.

In this section, we propose a set of axioms that describe the minimal requirements we wish to impose on a voter’s preferences over sets of pre-winners. The intuition underlying these axioms is that they would be satisfied by the preferences of any *rational* agent, irrespective of both the specific tie-breaking rule in use and any quantitative utilities the agent may assign to candidates. Our axioms do not define a total ordering over sets of pre-winners; they just specify conditions that any such ordering would satisfy. At the end of this section, we then give some examples for specific choices of preferences over sets of pre-winners, induced by specific choices for the tie-breaking rule and a voter’s expectations.

#### 3.1 General Axioms

We first present the axioms themselves and then discuss the underlying intuitions. Let  $\preceq$  be a voter’s preference ordering over candidates. We require the corresponding *preference ordering*  $\trianglelefteq$  over non-empty sets of pre-winners to satisfy the following five axioms ( $A$  and  $B$  denote non-empty sets):<sup>5</sup>

<sup>5</sup>Similar axioms are also given by Brams and Fishburn (1978). We note that there are conceivable rules for tie-breaking that do not satisfy our axioms (an example will

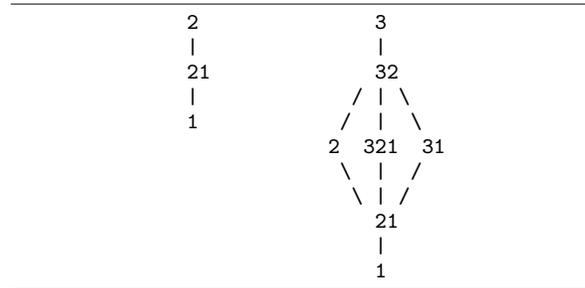


Figure 1: Preference orderings over pre-winners

- (REF)  $\trianglelefteq$  is reflexive.
- (TRA)  $\trianglelefteq$  is transitive.
- (DOM)  $A \trianglelefteq B$  holds whenever  $\#A = \#B$  and there exists a surjective mapping  $f : A \rightarrow B$  such that  $a \preceq f(a)$  for all  $a \in A$ .
- (ADD)  $A \trianglelefteq B$  holds whenever  $A \subset B$  and  $a \preceq b$  for all  $a \in A$  and all  $b \in B \setminus A$ .
- (REM)  $A \trianglelefteq B$  holds whenever  $B \subset A$  and  $a \preceq b$  for all  $a \in A \setminus B$  and all  $b \in B$ .

Axioms (REF) and (TRA) imply that  $\trianglelefteq$  is a *preorder*. These are very basic requirements for a rational preference ordering (Savage, 1972).

Axiom (DOM) may be interpreted as follows. If we are comparing two sets of candidates  $A$  and  $B$  and these sets have the same cardinality, then we can associate each candidate in  $A$  with a distinct candidate in  $B$  and make comparisons for these pairs. If there exists a pairing of candidates (given by  $f$ ) such that each candidate in  $B$  (weakly) *dominates* her counterpart in  $A$ , then also  $B$  as a whole should be preferred over  $A$ .

Axiom (ADD) can be read as expressing that, if we *add* a candidate  $b$  to a set of candidates  $A$  and  $b$  (weakly) dominates everyone in  $A$ , then the resulting set  $B$  should be preferred over  $A$  (the axiom says this for any number of candidates  $b$ , which boils down to the same thing). Analogously, axiom (REM) specifies that *removing* the least preferred candidate from a set should result in a new set that is preferred to the original. Note that we cannot make any general statements about the effects of adding or removing a candidate that is not extremal. For instance, it is unclear whether or not a set consisting of a voter’s favourite, second most favourite, and least preferred candidates (321) should be preferred over a set with just the second favourite candidate (2).

be given after Theorem 2), but we shall label these as “unreasonable” for the purposes of this paper.

For the cases of two and three pre-winners, respectively, the preorders generated by our axioms are shown in Figure 1. The case of four pre-winners is already somewhat difficult to draw. Nevertheless, using the axioms to check the relative preference of two given sets is easy. For example, we can verify that  $321 \trianglelefteq 42$  is bound to hold by first observing that  $321 \trianglelefteq 32$  follows from (REM), and then that  $32 \trianglelefteq 42$  is an instance of (DOM). The statement then follows from (TRA).

There is in fact a rich literature on ranking sets of objects given an existing ranking of the individual elements of such sets. Barberà et al. (2004) survey large parts of this literature. They distinguish three types of situations: (1) the elements of the set are interpreted as as possible outcomes after some randomised choice; (2) the elements of the set represent opportunities and the decision maker in question can choose one of them; or (3) the sets themselves are the final outcomes. Our axiom system falls under the first type (although certain refinements of the rankings induced by our axioms could also be seen as characterising opportunities).

### 3.2 Examples for Specific Choices

An example for a tie-breaking rule is *uniform tie-breaking*. This means that the winner is selected from the set of pre-winners such that each of the pre-winners has the same chance to get picked.

Next, we define two *refinements* of  $\trianglelefteq$  (as characterised by our axioms). The idea is that a specific voter would have a total preference ordering over sets of candidates, which satisfies the axioms, but which on top of that would also be guided by additional principles.

- Adopting the terminology of Taylor (2005), we say that a voter with preference relation  $\trianglelefteq$  is *optimistic* if he uses the following preference relation  $\trianglelefteq$  over sets of candidates:

$$A \trianglelefteq B \quad \text{iff} \quad \text{top}(A) \preceq \text{top}(B)$$

Here,  $\text{top}(C) \in \{c^* \in C \mid c \preceq c^* \text{ for all } c \in C\}$ . That is, the voter's preferences only depend on the *top candidate* within a set.<sup>6</sup> For example, we would get  $2 \trianglelefteq 31$  for an optimistic voter.

- In general, we only assume that a voter has an *ordinal* preference relation  $\preceq$ . But such an ordering *could* also be induced by a utility function  $u$  mapping candidates to numerical values. If this is the case and if a tie-breaking rule is used that assigns known probabilities to each event of picking

<sup>6</sup>This choice for  $\trianglelefteq$  is equivalent to what is called the *indirect-utility ranking of opportunity sets* in the survey paper by Barberà et al. (2004).

a particular winner from a given set of pre-winners (as is the case, for instance, for the uniform tie-breaking rule), then we may speak of voters that are *expected-utility maximisers*. For instance, if our voter assigns utilities  $u(3) = 10$ ,  $u(2) = 8$ ,  $u(1) = 2$  to the respective candidates, and if he is an expected-utility maximiser, then he would prefer 2 (expected utility 8) over 31 (expected utility 6) under uniform tie-breaking (*i.e.*  $31 \trianglelefteq 2$ ). Hence, this is an example where expected-utility maximisation under uniform tie-breaking and optimism give different results.

## 4 MANIPULATION UNDER APPROVAL VOTING

In this section we are going to analyse to what extent approval voting is prone to manipulation by individual voters. We assume that a voter will submit an insincere ballot only when doing so has the chance of securing a better outcome of the election for him than submitting any of the sincere ballots corresponding to his true preferences. That is, our notion of manipulability is somewhat more restrictive than the classical use of the term: a particular sincere ballot may well be dominated by an insincere ballot, but we do not consider this a problem as long as there is a further sincere ballot which in turn (weakly) dominates that insincere ballot. Also, we are not concerned with strategic considerations pertaining to the choice between alternative sincere ballots.

We begin by giving an example that shows that approval voting *is* manipulable in the above sense. Given the Gibbard-Satterthwaite Theorem (even when, as argued in the introduction, it does not immediately apply to our precise setting), this should not come as a surprise. However, some of the results presented in subsequent subsections may well be considered surprising. We are going to show that agents do not have an incentive to vote insincerely for four important and large classes of voting scenarios.

### 4.1 An Example of Successful Manipulation

Consider the case of four candidates. Suppose all but one voter have already cast their ballot (this is just another way of saying that the final voter *knows* how everyone else is going to vote). As before, let 4 be the candidate most preferred by this final voter, 3 is his second choice, 2 is his third choice, and 1 is his least preferred candidate. Suppose that 3 and 1 have each received 10 approvals, while 4 and 2 have only received 9 each. So our voter has the following opportunities (before tie-breaking):

- Force outcome 431 by voting with ballot [4].
- Force 3 by voting [43], [432], [3], or [32].
- Force 31 by voting [4321], [431], [321], or [31].
- Force 4321 by voting [42].
- Force 1 by voting [421], [41], [21], or [1].
- Force 321 by voting [2].

From the above, 431, 3, and 4321 are the only *undominated* sets of pre-winners. That is, these are the only outcomes that the voter can enforce for which there is no other outcome that is *definitely* preferred according to our five axioms for  $\preceq$ . Outcomes 431 and 3 can be enforced by means of sincere ballots (3 can also be enforced by means of the insincere ballots [3] and [32], but if this is the most preferred outcome for our voter, then there is no *need* to vote insincerely). 4321 is the only undominated outcome that can only be enforced by means of an insincere ballot (namely [42]). So our voter has an incentive to manipulate iff he strictly prefers 4321 over both 431 and 3. In all other cases, voting sincerely will be just as beneficial as manipulating, or even better.

Our example illustrates a *negative* result. But just *how negative* is it? This will depend on two points: (1) are there (many) more such examples?; and (2) is there a reasonable refinement of  $\preceq$  to a concrete total preference ordering over sets of pre-winners for our final voter such that the case captured by the example actually occurs? We are going to return to these questions in Section 4.4.

## 4.2 The Case of Optimistic Voters

We first show that manipulation can be avoided if all voters are *optimistic* (in the sense of Section 3.2). Optimism may stem from the tie-breaking rule, from the agents' quantitative utilities underpinning their ordinal preferences, or from a combination thereof.

**Theorem 1 (Optimistic voters)** *In approval voting, suppose that all but one voter have cast their ballot. Then, if the final voter is optimistic, he has no incentive to cast an insincere ballot.*

The proof of Theorem 1 can be found in the Appendix, at the end of this paper. Observe that Theorem 1 will apply in several (more or less) natural settings, e.g.:

- A voter will be optimistic if the *uniform tie-breaking rule* is used, and he is an *expected-utility maximiser* with *exponential preferences*. The latter means that he assigns quantitative utilities to

candidates  $c$  as follows:  $u(c) = m^{\#\{c' | c' \prec c\}}$ . For four candidates ( $m = 4$ ), for instance, he would assign utility 64 to his top choice, then 16, then 4, and finally 1 to his least preferred candidate. In combination with uniform tie-breaking, the *relative* expected utility for a set of pre-winners will then depend solely on the most preferred candidate in the set. For example, the expected utility of 421 would be  $E(421) = \frac{1}{3} \cdot (64 + 4 + 1) = 23$ , while that of 3 would be only  $E(3) = 16$ .

- Optimism may also stem from the tie-breaking rule alone. For instance, the voter for whom we want to ensure absence of incentives to manipulate may know (or believe) that the election chair will always break ties in his favour (Taylor, 2005).

It is not difficult to show that *pessimistic* voters (voters who assume that ties will always be broken *against* them) also do not have any incentive to vote insincerely: A final voter who is pessimistic should always vote for their favourite amongst the candidates receiving a maximal number of approvals from the other voters. Also voting for all the candidates ranked above that favourite top candidate (*i.e.* making the ballot sincere) will not affect the outcome.

## 4.3 The Case of Three Candidates

We now turn to a different kind of restriction. Instead on making additional assumptions on the voters' preferences over sets of candidates, we consider cases with a limited number of candidates. Recall that the Gibbard-Satterthwaite Theorem (like Arrow's Theorem) starts to *bite* when we move from two alternatives to three alternatives. As it turns out, maybe rather surprisingly, in our setting this transition only takes place when we move from three to four candidates. The example given in Section 4.1 shows that manipulation *can* occur in the case of four candidates, while our next theorem shows that manipulation never pays off if there are only three candidates. This theorem is a special case of a result by Brams and Fishburn (1978).<sup>7</sup> We are still going to give a direct proof here, because it demonstrates very well our approach taken later on in this paper.

**Theorem 2 (Three candidates)** *In approval voting with (at most) three candidates, suppose that all but one voter have cast their ballot. Then the final voter has no incentive to cast an insincere ballot.*

<sup>7</sup>The cited result concerns approval voting with voters with *trichotomous* preferences, *i.e.* (not necessarily strict) preference orderings that divide the set of candidates into (at most) three equivalence classes.

Situation	Sincere Ballots			Insincere Ballots			
	[3]	[32]	[321]	[31]	[2]	[21]	[1]
$\hat{3}\hat{2}\hat{1}$	<u>3</u>	32	321	31	2	21	1
$\hat{3}\hat{2}\dot{1}$	<u>3</u>	32	32	<u>3</u>	2	2	321
$\hat{3}\hat{2}\hat{1}$	<u>3</u>	32	32	<u>3</u>	2	2	32
$\hat{3}\hat{2}\dot{1}$	<u>3</u>	<u>3</u>	31	31	321	1	1
$\hat{3}\hat{2}\hat{1}$	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	32	321	31
$\hat{3}\hat{2}\hat{1}$	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	32	32	3
$\hat{3}\hat{2}\hat{1}$	<u>3</u>	<u>3</u>	31	31	31	1	1
$\hat{3}\hat{2}\dot{1}$	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	31	31
$\hat{3}\hat{2}\hat{1}$	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>	<u>3</u>
$\hat{3}\hat{2}\hat{1}$	<u>321</u>	<u>2</u>	21	1	<u>2</u>	21	1
$\hat{3}\hat{2}\hat{1}$	<u>32</u>	2	2	321	2	2	21
$\hat{3}\hat{2}\hat{1}$	<u>32</u>	2	2	<u>32</u>	2	2	2
$\hat{3}\hat{2}\hat{1}$	<u>31</u>	<u>321</u>	1	1	21	1	1
$\hat{3}\hat{2}\hat{1}$	<u>31</u>	<u>31</u>	1	1	1	1	1
$\hat{3}\hat{2}\hat{1}$	21	<u>2</u>	21	1	<u>2</u>	21	1
$\hat{3}\hat{2}\hat{1}$	<u>2</u>	<u>2</u>	<u>2</u>	21	<u>2</u>	<u>2</u>	21
$\hat{3}\hat{2}\hat{1}$	<u>2</u>	<u>2</u>	<u>2</u>	<u>2</u>	<u>2</u>	<u>2</u>	<u>2</u>
$\hat{3}\hat{2}\hat{1}$	1	<u>21</u>	1	1	<u>21</u>	1	1
$\hat{3}\hat{2}\hat{1}$	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>	<u>1</u>

Table 1: Enumeration of all relevant cases for Theorem 2

*Proof.* Trivial for less than three candidates. For three candidates, we have written a small program (in Prolog) that generates all relevant cases. The result is shown in Table 1. There is one row for every relevant “situation”. A situation is characterised by specifying which candidates are *pivotal* (having received a maximum number of approvals, before the final voter casts his ballot), which are *subpivotal* (having received exactly 1 approval less than the pivotal candidates), and which are *insignificant* (having received strictly fewer approvals than the subpivotal candidates, and thereby not having any chance of winning).<sup>8</sup> In the leftmost column, pivotal candidates are marked with a *hat* ( $\hat{\cdot}$ ), subpivotal candidates are marked with a *dot* ( $\dot{\cdot}$ ), while insignificant candidates are shown without any marking. For example,  $\hat{3}\hat{2}\hat{1}$  means that the final voter’s most preferred candidate (3) has no chance of winning, and that the other two candidates are pivotal.

The table has 7 columns, one for each possible ballot. Sincere ballots are grouped together on the left; insincere ballots are shown on the right. Inside the table, for each combination of a situation and a final ballot, we show the resulting set of pre-winners. To improve readability, *undominated* sets are underlined. For example, in situation  $\hat{3}\hat{2}\hat{1}$  (in the middle of the table), either 321 or 2 will beat any other feasible outcome.

<sup>8</sup>Table 1 has 19 rows as there are 19 relevant situations:  $3^3$  (choosing between *pivotal*, *subpivotal*, and *insignificant* for each of the 3 candidates), less  $2^3$  (to exclude those triples where no candidate is marked as *pivotal*).

While outcome 2 can be forced by the insincere ballot [2], it is also attainable by voting sincerely using the ballot [32]. If we check the entire table, we see that there is *no* situation where there would be an outcome that can be forced by an insincere ballot for which there is no sincere ballot achieving either the same or a better result. This proves the theorem.  $\square$

We stress that Theorem 2 is conditional on our axioms from Section 3.1. There are examples for tie-breaking rules that would breach these axioms. Suppose, for instance, that the election chair will break a tie between 1 and 2 in favour of 2, but if both 1 and 3 are amongst the pre-winners, she will always select 1 as the winner. This would result in  $31 \triangleleft 21$  and  $321 \triangleleft 21$ . Then, in situation  $\hat{3}\hat{2}\hat{1}$ , the final voter would have an incentive to submit the insincere ballot [21] (see Table 1).

#### 4.4 The Case of Four Candidates

If there are four candidates, the example given in Section 4.1 shows that manipulation is a possibility. Hence, we will not be able to generalise Theorem 2 to the case of four candidates. Instead, we now want to analyse *how many* cases there are where manipulation may occur.

To this end we have employed the same program used to generate the data for Table 1 to generate a large table showing the outcome for every combination of a “situation” (relative numbers of approvals before the

Sit.	Sincere Ballots				Insincere Ballots										
	[4]	[43]	[432]	[4321]	[431]	[42]	[421]	[41]	[3]	[32]	[321]	[31]	[2]	[21]	[1]
<u>4321</u>	<u>3</u>	32	321	31	2	21	1	<u>3</u>	32	321	31	2	21	1	
<u>432</u>	<u>3</u>	32	32	<u>3</u>	2	2	4321	<u>3</u>	32	32	<u>3</u>	2	2	321	
<u>432</u>	<u>3</u>	32	32	<u>3</u>	2	2	<u>432</u>	<u>3</u>	32	32	<u>3</u>	2	2	32	
<u>431</u>	<u>3</u>	<u>3</u>	31	31	<u>4321</u>	1	1	<u>3</u>	<u>3</u>	31	31	321	1	1	
<u>43</u>	3	3	3	3	432	4321	431	3	3	3	3	32	321	31	

Table 2: Five out of the 65 cases to consider for four candidates

final voter’s ballot is received) and a possible ballot for the final voter. This table has 15 columns (number of possible ballots for four candidates) and 65 ( $= 3^4 - 2^4$ ) rows (number of possible situations). We have then parsed this table for situations in which our voter may benefit from manipulation. The first 27 rows were particularly easy to check. These are all the rows with a  $\hat{4}$ , *i.e.* when candidate 4 is (one of) the pivotal candidate(s). In such a case, the sincere ballot [4] will always be optimal for the final voter and there is no incentive to manipulate. For illustration, we show the next five rows in Table 2. Again, outcomes that are definitely undominated are underlined. Of the cases shown, we can easily check that row 4 is the only example where the final voter has an incentive to manipulate: 4321 is neither dominated by 431 nor by 3 (this is actually the example familiar from Section 4.1). Now, the truly astonishing thing is that this is in fact *the only one of the 65 cases where an insincere ballot results in an undominated outcome!* In other words, we obtain the following theorem:<sup>9</sup>

**Theorem 3 (Four candidates)** *In approval voting with four candidates, suppose that all but one voter have cast their ballot. Then the final voter has no incentive to cast an insincere ballot, unless he strictly prefers 4321 over both 431 and 3.*

To appreciate the power of Theorem 3, consider this corollary (see Appendix for a detailed proof):

**Corollary 1 (Four candidates)** *In approval voting with four candidates and uniform tie-breaking, suppose that all but one voter have cast their ballot. Then, if*

<sup>9</sup>The full data and the Prolog program generating it are available on request. Besides producing the tables, the same program can also be used to automatically check this kind of data and to list all conditions on a voter’s preference ordering over sets of pre-winners that would allow for a situation where not voting sincerely could be the optimal strategy. This number of “exceptions” turns out to be fast increasing as the number of candidates increases. As we have seen, there are no exceptions for three candidates (Theorem 2) and just a single exception for four candidates (Theorem 3). For five candidates, there are 10 exceptions; for six candidates there are 63; for seven candidates there are 321; and for eight candidates there are 1447.

*the final voter is an expected-utility maximiser, he has no incentive to cast an insincere ballot.*

We stress that this last result does not rely on any assumptions regarding the actual utility values assigned to the different candidates by the final voter.

At this point, a few remarks regarding our general approach are in order. It may be argued that the assumption that our would-be manipulator knows exactly how many approvals each candidate is going to receive from everybody else is unrealistic. While this may be so, this actually strengthens our results: if you cannot gain any advantage from voting insincerely *even* when you have complete information, then that certainly speaks in favour of the voting rule. Still, if all that is known about the voting behaviour of the others are *mixed strategies* over submitting different ballots, then it may well be the case that voting sincerely is not in the best interest of our voter. This kind of model has been studied, amongst others, by De Sinopoli et al. (2006). For instance, suppose there are four candidates, we use approval voting with uniform tie-breaking, and all voters are expected-utility maximisers (*i.e.* as in Corollary 1). Now suppose there is just one other voter besides the would-be manipulator, and that other voter is using the following mixed strategy: with 50% probability submit ballot [43] and with 50% probability submit ballot [21]. Then it would be in the best interest of the final voter to use the (insincere) ballot [42], to ensure that for either one of the two possible cases his favourite amongst the two pivotal candidates will get elected.

#### 4.5 The Case of Expected-Utility Maximising Voters

We are now going to generalise the result presented in Corollary 1 and show that, under the assumption that all voters are *expected-utility maximisers*, no voter will have an incentive to vote insincerely under approval voting with *uniform tie-breaking*. In particular, like Theorem 1, this result is independent from the number of candidates running in an election. The proof may be found in the Appendix.

**Theorem 4 (Expected-utility maximisers)** *In approval voting with uniform tie-breaking, suppose that all but one voter have cast their ballot. Then, if the final voter is an expected-utility maximiser, he has no incentive to cast an insincere ballot.*

We believe that this is a result of particular practical relevance: uniform tie-breaking is intuitively a fair and acceptable tie-breaking rule that can be implemented in practice, and modelling voters as expected-utility maximisers is a standard assumption that aligns well with perceived reality.

## 5 CONCLUSIONS

We have started out with the observation that the classical impossibility result in voting theory, the Gibbard-Satterthwaite Theorem, *may* fail to apply in situations where there is more than one way of submitting a sincere ballot. Approval voting is such a system allowing for multiple sincere ballots. Our results have the flavour of results on strategy-proofness. Strategy-proofness in the classical sense, and in the context of a voting rule with *unique* sincere ballots, means that casting a sincere ballot is a (weakly) dominant strategy in all situations. Our theorems establish what could be termed  $\forall\exists$ -*strategy-proofness*: for any insincere ballot there exists a sincere ballot dominating the former.<sup>10</sup>

We have proved four such theorems for approval voting. The first makes certain assumptions on tie-breaking and/or voter preferences (optimism). The second puts a restriction on the number of candidates (it only applies to the case of three candidates), but is entirely general otherwise, without any assumptions regarding either tie-breaking or voter preferences. The third lies somewhere in between the first two: it applies

<sup>10</sup>Brams and Fishburn (1978) use the term *sincerity* (of a voting rule) to describe the same idea. A stronger notion would be  $\forall\forall$ -*strategy-proofness*: in any situation, all insincere ballots are dominated by all sincere ballots. For unique sincere ballots both versions coincide, with each other and with strategy-proofness in the classical sense. Another interesting concept to consider would be a notion of strategy-proofness that requires, for any situation and any insincere ballot  $X$ , the existence of a sincere ballot  $Y$  that dominates  $X$  and that *stands in a particular relationship* to  $X$ . For instance, we may require  $Y$  to be computable, given  $X$ , within a certain complexity class. (This may also raise new questions related to the flourishing work on computational barriers against manipulation (e.g. Conitzer et al., 2003; Elkind and Lipmaa, 2005): rather than asking how hard it is to manipulate a voting protocol, we may ask how much *harder* it is to identify a dominating sincere ballot as opposed to a dominating insincere ballot.) Or we may require some simple relationships between the “patterns” of the two ballots (e.g.  $Y$  may have to be the result of filling in all the gaps in  $X$ ; or  $X$  and  $Y$  may have to make the same number of approvals).

to four candidates, and it only puts a very mild restriction on voter preferences (excluding just one specific scenario). Importantly, these mild restrictions cover the case that appears to be both the most natural and the the most relevant in practice, namely when ties are broken using a uniform probability distribution and voters can be assumed to aim at maximising their expected utility. Finally, our fourth and final theorem generalises this observation and shows that under uniform tie-breaking and given the assumption of expected-utility maximisation on behalf of the voters, there will be no incentive to vote insincerely — whatever the number of candidates in the race.

We believe that these results reveal some attractive aspects of approval voting, which appear to be relevant in practice and which may have been previously overlooked due to the overpowering presence of seemingly all-encompassing impossibility results. However, we should also stress that our results should not be misinterpreted as saying that approval voting would not provide voters with an incentive to strategise at all. Our results only show that, when considering what ballot to submit, voters can concentrate on the set of sincere options at their disposal, as none of the (exponential number of) insincere ballots can ever yield a strictly better outcome for them (assuming that the side conditions of one of our theorems apply). Strategising *within* the (much smaller) space of sincere ballots is still possible and often beneficial (as is apparent from Table 1, for instance).

Some of the results in this paper have been derived with the assistance of a computer. More specifically, we have used a Prolog program to generate all relevant cases and to present them in a form that makes them easy to check. We have also reported on the use of a version of this program which also automates the checking phase. It is likely that this approach can be extended to other settings as well. For instance, one could investigate what happens if we change the set of axioms defining voter preferences over sets of candidates before tie-breaking, or to what extent groups of voters (rather than individuals) may benefit from voting insincerely. Finally, we have started the paper by pointing out the fact that the presence of multiple sincere ballots may allow us to (in some sense) circumvent the Gibbard-Satterthwaite Theorem. This observation is independent from the specific framework of approval voting, and it may be possible to apply our approach also in other settings.

## Acknowledgements

I would like to thank Steven Brams, Peter van Emde Boas, Eric Pacuit, Olivier Roy and Jelle Zuidema for helpful discussions and feedback on an earlier ver-

sion of this paper. I would also like to thank the three anonymous TARK reviewers for their excellent reviews. One of them in particular conjectured something equivalent to Theorem 4 (which I had only worked out after the submission deadline).

## References

- S. Barberà, W. Bossert, and P. Pattanaik. Ranking sets of objects. In *Handbook of Utility Theory: Volume 2*. Kluwer Academic Publishers, 2004.
- S. J. Brams and P. C. Fishburn. Approval voting. *American Political Science Review*, 72(3):831–847, 1978.
- V. Conitzer, J. Lang, and T. Sandholm. How many candidates are required to make an election hard to manipulate? In *Proc. TARK-2003*, 2003.
- F. De Sinopoli, B. Dutta, and J.-F. Laslier. Approval votig: Three examples. *International Journal of Game Theory*, 35:27–38, 2006.
- E. Elkind and H. Lipmaa. Hybrid voting protocols and hardness of manipulation. In *Proc. 16th International Symposium on Algorithms and Computation (ISAAC-2005)*. Springer-Verlag, 2005.
- A. Gibbard. Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601, 1973.
- R. G. Niemi. The problem of strategic behavior under approval voting. *American Political Science Review*, 78(4):952–958, 1984.
- R. Parikh and E. Pacuit. Safe votes, sincere votes, and strategizing. Presented at the 16th International Conference on Game Theory at Stony Brook, 2005.
- M. A. Satterthwaite. Strategy-proofness and Arrow’s conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217, 1975.
- L. J. Savage. *The Foundations of Statistics*. Dover Publications, 2nd edition, 1972.
- A. D. Taylor. *Social Choice and the Mathematics of Manipulation*. Cambridge University Press, 2005.

## APPENDIX

### Proof of Theorem 1

Call the final voter  $\ell$ . Let  $C^{++}$  be the set of candidates that have received the highest number of approvals from the other voters (these are the *pivotal* candidates), and let  $C^+$  be the set of candidates that have each received exactly 1 approval less than those in  $C^{++}$  (these are the *subpivotal* candidates). Only

pivotal and subpivotal candidates have a chance of winning: the set of pre-winners will be a subset of  $C^{++} \cup C^+$ . Any approvals that  $\ell$  may cast for other candidates will be irrelevant. There are two possible scenarios:

- (1) The set of pre-winners is a subset of  $C^{++}$  (namely, if  $\ell$  approves of that subset and any number of candidates outside of  $C^{++}$ ).
- (2) The set of pre-winners is the union of a subset of  $C^+$  and the full set  $C^{++}$  (namely, if  $\ell$  approves of that subset of  $C^+$  and any number of candidates outside of  $C^{++} \cup C^+$ ).

From amongst the cases covered by scenario (1), the best outcome for  $\ell$  is when only his most preferred candidate in  $C^{++}$ , let’s call her  $top_\ell(C^{++})$ , is a pre-winner. This follows from axiom (REM). Voter  $\ell$  can force this preferable outcome by simply approving of  $top_\ell(C^{++})$  alone. Furthermore,  $\ell$  can turn that ballot into a sincere ballot by also approving of everyone he rates above  $top_\ell(C^{++})$ , which will not change the outcome. Hence, for scenario (1) we are done (note that, so far, we did not have to refer to  $\ell$  being optimistic).

Now suppose we are in scenario (2). If the preferred pre-winner is in  $C^{++}$ , *i.e.* if  $top_\ell(C^{++} \cup C^+) = top_\ell(C^{++})$ , then  $\ell$  can secure a preferable outcome by just approving of  $top_\ell(C^{++})$ . This is the same argument as above. Also, making that ballot sincere by also approving of everyone higher up in the ranking will not alter the result. Hence,  $\ell$  has no incentive to cast an insincere ballot.

So the final case to consider is scenario (2), with the assumption that  $top_\ell(C^{++} \cup C^+) = top_\ell(C^+)$ . (This will be the only point where we require  $\ell$  to be optimistic.) To get into scenario (2), voter  $\ell$  had to approve of a subset of  $C^+$ . Now suppose he approves of just  $top_\ell(C^+)$  instead. The new outcome would be  $C^{++} \cup \{top_\ell(C^+)\}$ . Optimism allows us to remove elements from a set of pre-winners without affecting the ranking of that set, as long as we leave the top pre-winner in place. Hence, the new outcome is as preferred as the original one. Turning the ballot of approving of  $top_\ell(C^+)$  alone into a sincere ballot (by approving of everyone above as well) will either not change the outcome at all, or result in a subset of  $C^{++}$ , all of which are ranked above  $top_\ell(C^+)$  by  $\ell$ , becoming the sole pre-winners. Either way, voter  $\ell$  has no incentive not to submit a sincere ballot.  $\square$

### Proof of Corollary 1

By Theorem 3, the claim will certainly hold if we can be sure that the final voter does not rate 4321

strictly above both 431 and 3. Let  $u_4$  be the utility that the voter assigns to candidate 4, and similarly for  $u_3, u_2, u_1$ . Note that we have:

$$u_4 \geq u_3 \geq u_2 \geq u_1 \quad (1)$$

If the voter is an expected-utility maximiser and if we use uniform tie-breaking, then the expected utilities of the three outcomes under consideration are as follows:

$$E(4321) = \frac{1}{4} \cdot (u_4 + u_3 + u_2 + u_1) \quad (2)$$

$$E(431) = \frac{1}{3} \cdot (u_4 + u_3 + u_1) \quad (3)$$

$$E(3) = u_3 \quad (4)$$

Now assume that our voter strictly prefers 4321 over 431 and 3. This means that  $E(4321) > E(431)$  and  $E(4321) > E(3)$ . Hence, we also get the following inequalities:

$$\frac{1}{4} \cdot (u_4 + u_3 + u_2 + u_1) > \frac{1}{3} \cdot (u_4 + u_3 + u_1) \quad (5)$$

$$\frac{1}{4} \cdot (u_4 + u_3 + u_2 + u_1) > u_3 \quad (6)$$

These last two inequalities can be rewritten as follows:

$$3 \cdot u_2 > u_4 + u_3 + u_1 \quad (7)$$

$$u_4 + u_2 + u_1 > 3 \cdot u_3 \quad (8)$$

But together with (1), these two are inconsistent. Hence, we have derived a contradiction and shown that the only critical case left open by Theorem 3 can in fact never occur under uniform tie-breaking with expected-utility maximising voters.  $\square$

### Proof of Theorem 4

Recall the concepts of pivotal, subpivotal, and insignificant candidates (as introduced in the proof of Theorem 2). Either the final voter has or has not approved of a pivotal candidate. In the former case, the situation is the same as in the proof of Theorem 1: The approved pivotal candidates are the pre-winners, and we can do at least as well by approving only of the most preferred pivotal candidate. This ballot in turn yields the same outcome as the (sincere) ballot approving of the most preferred pivotal candidate together with any (subpivotal or insignificant) candidates rated above that candidate.

Now suppose our voter has not approved of any of the pivotal candidates. Then the set of pre-winners will be the set of pivotal candidates together with the set of approved subpivotal candidates. We are going to describe a procedure for turning an insincere ballot of this kind into a sincere ballot that will increase (or at least maintain) the expected utility of the outcome. Let  $c$  be a disapproved candidate that is preferred over at least one of the approved candidates (such a candidate  $c$  exists, otherwise the ballot would be sincere

already). Our procedure for turning an insincere ballot into a sincere ballot consists of repeatedly applying the following transformation step:

- Suppose that  $c$  is *insignificant*. In this case we simply approve of  $c$  in the new ballot (this does not change the outcome).
- Now suppose that  $c$  is *subpivotal*. We distinguish the following two cases:
  - There is a subpivotal candidate  $c'$  that is less preferred than  $c$  and that our voter has approved of. In this case we change the ballot by approving of  $c$  and disapproving of  $c'$ . This means that  $c'$  will be replaced by  $c$  in the set of pre-winners, which will increase expected utility.
  - There is no such subpivotal candidate, *i.e.* all approved candidates that are less preferred than  $c$  are insignificant. In this case, simply disapprove of all of these insignificant candidates (which does not affect the outcome).
- Finally, suppose that  $c$  is *pivotal*. We again distinguish two cases:
  - The utility of  $c$  is at least as high as the average utility of the pre-winners. In this case, we change the ballot completely: *only* approving of  $c$  (pivotal) will make  $c$  the unique winner and hence increase (or maintain) expected utility. As argued earlier, this new ballot is in turn weakly dominated by a vote for the most preferred pivotal candidate (together with any preferred insignificant or subpivotal candidates, which may be added to make the ballot sincere).
  - The utility of  $c$  is lower than the average utility of the pre-winners. In this case, losing any of the candidates from the set of pre-winners that are less preferred than  $c$  will increase expected utility. So in the new ballot, we can disapprove of all subpivotal (and insignificant) candidates that are less preferred than  $c$  (and keep  $c$  disapproved of) and increase expected utility in the process.

Clearly, this procedure must terminate (eventually, there will be no more candidate  $c$  violating sincerity). As each single transformation either increases or maintains expected utility, the final ballot results in an outcome that (weakly) dominates the outcome forced by the original (insincere) ballot. Hence, our voter can always do at least as well by voting sincerely as by voting insincerely.  $\square$