How Does Uncertainty about Other Voters Determine a Strategic Vote?

Zeinab Bakhtiari   Hans van Ditmarsch   Abdallah Saffidine

Abstract.   We propose a framework for strategic voting when a voter may lack knowledge about the preferences of other voters, or about other voters' knowledge about her own preference. In this setting we define notions of manipulation, equilibrium, and dominance, under uncertainty. We provide scenarios wherein the profiles of true preferences are the same but the equilibrium profiles are different, because the voters have different knowledge about other voters. We also model actions that change such uncertainty about preferences, such as a voter revealing her preference. We show that some forms of manipulation and equilibrium are preserved under such uncertainty updates and others not. We then formalize epistemic voting terminology in a logic. Our aim is to provide the epistemic background for the analysis and design of voting rules that incorporate uncertainty.

1 Introduction

A well-known fact in social choice theory is that strategic voting, also known as manipulation, becomes harder when voters know less about the preferences of other voters. Standard approaches to manipulation in social choice theory [17, 28] as well as in computational social choice [6] assume that the manipulating voter knows the sincere preferences of other voters. Some approaches [15, 5] assume that voters have a probabilistic prior belief on the outcome of the vote, which encompasses the case where each voter has a probability distribution over the set of profiles. In yet other approaches the uncertainty of the manipulator is modelled as the inability to distinguish between a set of voting profiles [12, 23]. In [12] manipulators have incomplete knowledge of the non-manipulators' preferences. Meir ([23]) provides a setting for iterated voting, wherein incomplete knowledge of the profile at the next iteration is induced by the partial view on the profile at the current iteration. Still, we think that the study of strategic voting under complex belief states has received little attention so far, especially when voters are uncertain about the uncertainties of other voters, i.e., when we model higher-order beliefs of voters.

In this contribution we present a logic to model higher-order uncertainty of voters. On the assumption that voters may be uncertain about other voters' preferences but know their own preference, we model how this uncertainty may determine a strategic vote, and how a reduction in this uncertainty may change a strategic vote. We give scenarios where the profile is the same, and even the set of profiles about which the manipulator is uncertain is the same, but where the uncertainty about other voters is different, thus resulting in different manipulations. Additionally, reducing such uncertainty may affect manipulative behaviour, and we also give example scenarios for that. We model uncertainty reducing actions as truthful public announcements [26].

There are several ways of expressing incomplete knowledge about the linear order of preference of a voter. The literature on possible and necessary winners assumes that it is expressed by a collection of partial strict orders (one for each voter), while [19] consider it to consist of a collection of probability distributions, or a collection of sets of linear orders (one for each voter), i.e., a collection of profiles. For partial preference see also [21] and more generally for voting under incomplete knowledge see [9].

A first link between epistemic logic and voting, to our knowledge, has been given in [10]—they use knowledge graphs to indicate that a voter is uncertain about the preferences of another voter. A more recent follow-up of that, within the area known as social software, is [24].

An independent line of modal logics for social choice, of which voting can be seen as a special case, was proposed in [3] and in the journal follow-up [2]. They consider two modalities of which one formalizes what is true for all profiles (in the current agenda). It is clearly similar to our epistemic modality formalizing uncertainty about (voting) profiles. The quantification in [2] is game theoretical rather than epistemic as in [10] and in our proposal.

Modal logics of social choice and of voting have further been proposed in [29] and, building on that, in [11, 25]. Also in these logics the semantic primitives are preferences of agents or voters, and sets of those, i.e., profiles. But the modalities encode agency, and not uncertainty, as we do. Compared to their results the logical equivalents of our voting primitives are very 'flat': essentially we encode them as big disjunctions of profiles, for example, we represent the proposition that a is the winner of the election by the (very large) disjunction of all profiles wherein this is the case, given that the voting function F is a parameter of the logic (and not of the model, as in [11]). In our case, all the logical action goes into the uncertainty about profiles, and the modelling results involve the formalization of epistemic notions such as conditional (i.e., Bayesian) equilibrium, and their (epistemic) updates.

The uncertainty of a manipulating voter in the mentioned [12] and [23] is modelled in information sets (i.e., set of indistinguishable profiles). However, from that voter's perspective the other voters are not uncertain, so that higher-order uncertainty is not considered. The goals of [23] are similar to ours (when do equilibria exist,

assume risk aversion) but his methods are statistical (there is no higher-order uncertainty).

Our setting shares also some similarity with robust mechanism design [8], which generalizes classical mechanism design by weakening the common knowledge assumptions of the environment among the players and the planner. In [8] uncertainty is modelled with information partitions. The main technical difference is that in our setting, as in classical social choice theory, preferences are ordinal, whereas in (robust) mechanism design preferences are numerical payoffs, which allows for payments.

This is an overview of our contribution. Section 2 presents voting terminology. Section 3 introduces knowledge profiles, our semantic primitive. Section 4 investigates epistemic notions of dominance. Section 5 defines equilibrium profiles under uncertainty, wherein voting is represented as a Bayesian game. Section 6 is devoted to uncertainty updates and how this affects knowledge of other voters and equilibrium profiles. Then, in Section 7 we succinctly present the logic of this contribution more formally — we do this at the end, because our main focus is the semantic interaction of knowledge and voting, not the logic.

2   Voting

Assume a finite set $\mathcal{N} = \{1, \dots, n\}$ of n *voters* (or *agents*), and a finite set $\mathcal{C} = \{a, b, c, \dots\}$ of m *candidates* (or *alternatives*). Voter variables are $i, j, \dots$, and candidate variables are $x, y, (x_1, x_2,)\dots$. Let $O(\mathcal{C})$ be the set of linear orders on $\mathcal{C}$.

Definition 1 (Preference, profile, voting rule, vote)   For each voter i, a *preference (relation)* (over $\mathcal{C}$) is a linear order on $\mathcal{C}$. A *profile* $\succ$ is a function $\succ: \mathcal{N} \to O(\mathcal{C})$ assigning a preference $\succ(i)$, denoted as $\succ_i$, to each voter. A *(resolute) voting rule* is a function $F : O(\mathcal{C})^{\mathcal{N}} \to \mathcal{C}$ from the set of profiles for $\mathcal{N}$ to the set of candidates. From the perspective of the voting rule the enacted preference $\succ_i$ is a *vote*.

If voter i prefers candidate a to candidate b in preference $\succ_i$, we write $a \succ_i b$, or $b \prec_i a$. For ($a = b$ or $a \succ_i b$) we write $a \succeq_i b$, or $b \preceq_i a$. Profile variables are $\succ, \succ', \succ'', \dots$. If $\succ \in O(\mathcal{C})^{\mathcal{N}}$ and $\succ'_i \in O(\mathcal{C})$, then $(\succ_{-i}, \succ'_i)$ is the profile wherein $\succ_i$ is substituted by $\succ'_i$ in $\succ$.

The voting rule determines which candidate wins the election — $F(\succ)$ is the *winner*. In case there is more than one *tied co-winner* we assume a *tie-breaking preference*, that is a linear order over candidates. In the *plurality voting rule* the winner is the candidate who is most often ranked the top candidate (most preferred in $\succ_i$), where in case there are several co-winners the tie-breaking preference selects one.

Voters cannot be assumed to vote according to their preference. Instead of giving her *sincere* or *truthful preference*, a voter may cast another preference as her vote. This is an *insincere* or *strategic preference*. If that vote improves the outcome it is a *manipulation*.

Definition 2 (Manipulation)   Let $i \in \mathcal{N}$, $\succ \in O(\mathcal{C})^{\mathcal{N}}$, $\succ'_i \in O(\mathcal{C})$, and F a voting rule. If $F(\succ_{-i}, \succ'_i) \succ_i F(\succ)$, then $\succ'_i$ is a *manipulation* by voter i of profile $\succ$.

The combination of a profile $\succ$ and a voting rule F defines a strategic game: a player is a voter, an individual strategy for a player is a preference, a strategy profile (of players) is therefore a profile in our defined sense (of voters), and the preference of a player among the outcomes is according to her sincere vote: given profiles $\succ'$, $\succ''$, voter i also prefers outcome $F(\succ')$ over outcome $F(\succ'')$ in the game theoretical sense iff (in the voting sense) $F(\succ') \succ_i F(\succ'')$. The relevant notions of dominance and equilibrium are as follows.

Definition 3 (Dominant preference)   Let $i \in \mathcal{N}$, $\succ \in O(\mathcal{C})^{\mathcal{N}}$, $\succ'_i \in O(\mathcal{C})$, and F a voting rule. If for all $\succ'' \in O(\mathcal{C})^{\mathcal{N}}$, $F(\succ''_{-i}, \succ'_i) \succeq_i F(\succ'')$, and for some $\succ'' \in O(\mathcal{C})^{\mathcal{N}}$, $F(\succ''_{-i}, \succ'_i) \succ_i F(\succ'')$, then $\succ'_i$ is a *dominant preference* for voter i.

A dominant preference corresponds to a dominant strategy in game theory.[1] The reader will recognize this as *weak* dominance. We may also use *strong* dominance, which holds for $\succ'_i$ if for all $\succ'' \in O(\mathcal{C})^{\mathcal{N}}$, $F(\succ''_{-i}, \succ'_i) \succ_i F(\succ'')$. Note that a dominant preference need not be a manipulation of $\succ$: the strict part may be satisfied for another profile than the profile $\succ$ of true preferences. (Of course, a *strongly* dominant strategy is also a manipulation.)

Definition 4 (Equilibrium profile)   A profile $\succ$ is an *equilibrium profile* iff no voter has a manipulation of $\succ$.

An equilibrium profile corresponds to a Nash equilibrium (in game theory). An equivalent way of defining equilibrium profile is: A profile $\succ$ is an equilibrium profile iff for all $\succ'_i \in O(\mathcal{C})$, $F(\succ) \succeq_i F(\succ_{-i}, \succ'_i)$. The formulations correspond:

Suppose the above condition does not hold. Then there is a voter i and some preference $\succ'_i \in O(\mathcal{C})$ such that $F(\succ) \prec_i F(\succ_{-i}, \succ'_i)$, i.e., $F(\succ_{-i}, \succ'_i) \succ_i F(\succ)$, i.e., $\succ'_i$ is a manipulation for voter i of profile $\succ$. In other words, if the condition holds, then no voter has a manipulation of $\succ$.

3   Knowledge profiles

We model uncertainty in voting as incomplete knowledge about the profile. The structures to represent such uncertainty are standard in modal logic [16, 13]. To allow for the definition of dominance and of equilibria under uncertainty, we require that voters know their own preference.

---

[1]We called this (game theoretical) notion 'dominant preference' and not 'dominant manipulation', to avoid confusion with the (epistemic) notion of *dominant manipulation of an information set*, defined in Subsection 4.2: only the latter is commonly used in voting theory, as the requirements of the former are very strong.

Definition 5 (Knowledge profile)   A *profile model* is a structure $\mathcal{M} = (S, \sim, \pi)$, where S is a *domain* of abstract objects called states; where $\sim : \mathcal{N} \to \mathcal{P}(S \times S)$ is a function such that for $i = 1, \ldots, n$, ($\sim(i)$ written as) $\sim_i$ is an *indistinguishability relation* that is an equivalence relation; and where *valuation* $\pi : S \to O(\mathcal{C})^{\mathcal{N}}$ assigns a profile to each state and such that $s \sim_i t$ implies $\pi(s)_i = \pi(t)_i$. The *information set* of voter i in state s is defined as $[s]_{\sim_i} := \{t \mid s \sim_i t\}$. Let $\pi([s]_{\sim_i})$ denote $\{\pi(t) \mid s \sim_i t\}$. A *knowledge profile* is pointed structure $\mathcal{M}_s$ where $s \in S$.

Unless confusion results, the set $\pi([s]_{\sim_i})$ of profiles that i considers possible is also called an information set, as in the voting literature [12]. In general, different information sets (i.e., sets of states) may be about the same set of profiles.

Definition 6 (Knowledge and ignorance)   Given is a knowledge profile $\mathcal{M}_s$, where $\mathcal{M} = (S, \sim, \pi)$. Let φ be a proposition about profiles. Voter i *knows* φ in $\mathcal{M}_s$, iff φ is true in all $t \in S$ such that $s \sim_i t$. Voter i *considers possible that* (or *does not know that not*) φ in $\mathcal{M}_s$, iff φ is true in some $t \in S$ such that $s \sim_i t$; if, in that case, there is an additional state $u \in S$ with $s \sim_i u$ in which φ is false, then we say that i *does not know whether* (or *is uncertain about*, or *is ignorant about*) φ.
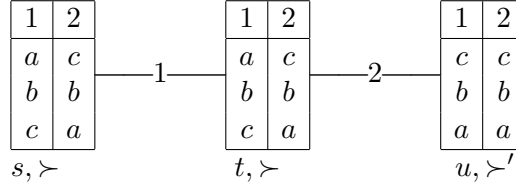
Section 7 contains a proper inductive definition of 'proposition about profiles', and a formal semantics. For now it suffices to say that the following are propositions about profiles: $\succ$ or 'the profile is $\succ$'; $a \succ_i b$, and $\succ_i$ ('the preference of voter i is $\succ_i$'); and also propositions like 'voter i knows that voter j knows $\succ_i$'. Example 7 demonstrates that: *different states may be assigned the same profile, but have different knowledge properties*. In scenarios where different states are always assigned different profiles, we can say that the uncertainty of a voter is (only) about a collection of profiles. But in scenarios where different states are assigned the same profile, the set $\pi([s]_{\sim_i})$ of profiles that voter i considers possible is smaller than the set $[s]_{\sim_i}$ of states that i considers possible.

Partial preferences cannot be expressed in our framework. In particular, uncertainty between $a \succ_i b \succ_i c$ and $b \succ_i a \succ_i c$ does not mean indifference between candidates a and b. Uncertainty between $a \succ_i b \succ_i c$ and $b \succ_i a \succ_i c$ means that ($a \succ_i b \succ_i c$ or $b \succ_i a \succ_i c$) is true. This entails ($a \succ_i b$ or $b \succ_i a$), which is equivalent to (not ($b \succeq_i a$ and $a \succeq_i b$)). That is the opposite of indifference between a and b, as that means ($b \succeq_i a$ and $a \succeq_i b$).

Example 7   Consider two voters who are Leela (1) and Sunil (2), children of Devi, and who 'vote' for an animated movie to see before bedtime; where the choice is between a (Alice in Wonderland), b (Brave), and c (Cars). Leela's preference is $a \succ_1 b \succ_1 c$ and Sunil's preference is $c \succ_2 b \succ_2 a$. Sunil (2) is uncertain if Alice is Leela's most or least preferred movie, and dually if Cars is Leela's least or most preferred movie. More interestingly, Leela (1) knows Sunil's preference, but she is uncertain whether Sunil knows her preference. Even more interestingly, Sunil is also uncertain

whether, in case Alice is Leela's most preferred movie, she knows that he does not know that.

We model this as a knowledge profile $\mathcal{M}_t$ consisting of three states s, t, u and for two voters 1 and 2. State s is assigned to profile $\succ$, wherein $a \succ_1 b \succ_1 c$ and $c \succ_2 b \succ_2 a$, etc. Below, a column represents a preference relation, and states that are indistinguishable for a voter i are linked with an i-labelled edge. The partition for 1 on the domain is therefore $\{\{s, t\}, \{u\}\}$, and the partition for 2 on the domain is $\{\{s\}, \{t, u\}\}$.

$$
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
a & c \\
b & b \\
c & a \\
\hline
\end{array}
\;-1-\;
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
a & c \\
b & b \\
c & a \\
\hline
\end{array}
\;-2-\;
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
c & c \\
b & b \\
a & a \\
\hline
\end{array}
$$
$$
s, \succ \qquad\qquad t, \succ \qquad\qquad u, \succ'
$$

States s and t have been assigned the same profile $\succ$ but have different epistemic properties. In s, 2 knows that 1 prefers a over c, whereas in t, 2 does not know that. We list some relevant propositions that are true in the actual state t:

- Leela prefers Alice over Cars. *This is true, because $a \succ_1 c$ in t.*
- Sunil does not know that Leela prefers Alice over Cars. *This is true, because $t \sim_2 u$, and $a \succ_1 c$ is false in u.*
- Leela knows Sunil's preference, but she is uncertain whether Sunil knows her preference. *In s, Sunil knows that Leela's preference is $a \succ_1 b \succ_1 c$, whereas in t, Sunil does not know that Leela's preference is $a \succ_1 b \succ_1 c$, because $t \sim_2 u$, and $c \succ_1 b \succ_1 a$ in u.*

Unlike merely sincere and insincere preference, in knowledge profiles there are *three* kinds of preference: actual sincere preference, possible sincere preference, and insincere preference. In $\mathcal{M}_t$ of Example 7, Leela's actual sincere preference is $a \succ_1 b \succ_1 c$, a possible sincere preference is $c \succ_1 b \succ_1 a$ (namely in state u, from the perspective of Sunil), and an insincere preference is $b \succ_1 a \succ_1 c$. This can be confusing.

4 Manipulation, knowledge and dominance

4.1 Manipulation and knowledge

Given are a knowledge profile $\mathcal{M}_s$ where $\pi(s) = \succ$, and a voting rule F. If voter i can manipulate $\succ$, then voter i can also manipulate $\mathcal{M}_s$. This is because manipulation is defined with respect to the profile of the actual state of the knowledge profile (it is a game theoretical notion, not an epistemic notion). So it may be that a voter can manipulate the vote but does not know that, because she considers another profile possible wherein she cannot manipulate the vote. Notions of manipulation that

involve knowledge can be borrowed from the knowledge and action literature [7, 20]. A curious situation is when in all states that the voter considers possible there is a manipulation, but when in different such states there are different manipulations. So she knows that she has a manipulation, but she does not know what the manipulation is. This is called *de dicto knowledge* of manipulation. A stronger form of knowledge is when there is a preference $\succ_i'$ that is *the same* manipulation in any state that the voter considers possible. This is called *de re knowledge* of manipulation.

Definition 8 (Knowledge of manipulation)   Given are a knowledge profile $\mathcal{M}_s$ and a voting rule F.

- Voter i knows *de dicto* that she can manipulate $\mathcal{M}_s$, if for all profiles $\succ \in \pi([s]_{\sim_i})$ there is a preference $\succ_i'$ such that $\succ_i'$ is a manipulation in $\succ$.
- Voter i knows *de re* that she can manipulate $\mathcal{M}_s$, if there is a preference $\succ_i'$ such that for all profiles $\succ \in \pi([s]_{\sim_i})$, $\succ_i'$ is a manipulation in $\succ$.

If voter i knows *de re* that she can manipulate the election, she has the ability to manipulate, namely by strategically voting $\succ_i'$. But in *de dicto* manipulations the voter does not seem to have that ability. It is akin to 'game of chicken' type equilibria in game theory, wherein for each strategy of a player there is a complementary strategy of the other player such that the pair is an equilibrium, but where this choice cannot be made without coordination. An example of de dicto knowledge of manipulation for Borda voting is given in [14].

Consider the profile model $\mathcal{H}$ consisting of the domain $O(\mathcal{C})^{\mathcal{N}}$, so we can identify states s with their profiles $\succ = \pi(s)$, and such that all voters only know their own preferences: $\succ \sim_i \succ'$ iff $\succ_i = \succ_i'$. In this (unique) model it is common knowledge that voters only know their own preferences. We can see it as an *interpreted system* [16] consisting of *global states* that are profiles and where *local states* are individual preferences. A model such as $\mathcal{H}$ is known as a *hypercube* [22].

Proposition 9   In the hypercube profile model, knowledge of manipulation is impossible for plurality voting.

Proof   The result holds for *de re* knowledge of manipulation and for *de dicto* knowledge of manipulation. We start with the *de re* case.

Let $\succ$ be the profile. Let us assume that there are a sufficient number of voters and candidates to avoid boundary cases. Assume towards a contradiction that voter i knows that $\succ_i'$ is a manipulation. As i is uncertain about the preferences of other voters, she considers it possible all other voters j have the same preferences as herself, i.e., she considers it possible that the profile is $\succ'' \in O(\mathcal{C})^{\mathcal{N}}$ such that for all $j \in \mathcal{N}$, $\succ_j'' = \succ_i$. In that case, i's preferred candidate would have won by majority vote, contradicting the assumption that $\succ_i'$ is a manipulation.

In the *de dicto* case, for each profile that voter i considers possible there is a

manipulation. For all those profiles, assuming that all other voters j have the same preference as i, again derives a contradiction. $\square$

4.2 Dominant manipulation and knowledge

We now compare the notions of manipulation and dominant strategy with the notion of *dominant manipulation of an information set* in voting theory [12].

Definition 10 (Dominant manipulation of an information set)　Let a knowledge profile $\mathcal{M}_s$ with $\pi(s) = \succ$, $i \in \mathcal{N}$, $\succ'_i \in O(\mathcal{C})$, and a voting rule F be given. If for all $\succ'' \in \pi([s]_{\sim_i})$, $F(\succ''_{-i}, \succ'_i) \succeq_i F(\succ'')$, and for some $\succ'' \in \pi([s]_{\sim_i})$, $F(\succ''_{-i}, \succ'_i) \succ_i F(\succ'')$, then $\succ'_i$ is a *dominant manipulation* for voter i of information set $\pi([s]_{\sim_i})$ (or: of knowledge profile $\mathcal{M}_s$).

Observe that dominant manipulation of an information set according to [12] is on the assumption that all other voters vote according to their true preference, whereas dominant preference in the game theoretical sense (Def. 3) is on the assumption that all other voters can choose any preference as their vote. The first is dominance **no matter the true preference** of others (but assuming that this is their vote), the second is dominance **no matter the vote** of others (but assuming what their true preferences are). We consider this difference curious. However, despite such seemingly orthogonal epistemic and game-theoretical dimensions, they are after all very much related, as now shown in the following proposition. We therefore find the observations made in this proposition, although elementary, somewhat surprising.

Proposition 11

1. A dominant manipulation of a singleton set $\{\succ\}$ is a manipulation of $\succ$ (Def. 2).
2. A dominant manipulation of the hypercube knowledge profile $\mathcal{H}_s$ is a dominant preference given profile $\pi(s)$ (Def. 3).
3. If a voter has a dominant manipulation then she *knows* that she has a dominant manipulation.
4. If a voter has a dominant manipulation then she may not have a manipulation.
5. Knowledge of dominant manipulation does not imply knowledge of manipulation (neither *de re* nor *de dicto*) (Def. 8).
6. If a voter has *de re* knowledge of manipulation then she has a dominant manipulation.

Proof

1. The strictness requirement of dominance must apply to $\succ$.
2. See the curious observation above. We recall that the notion of *dominant manipulation* models uncertainty over the true preferences of others (the 'epistemic dimension'), whereas the notion of *dominant preference* models uncer-

tainty over how others vote (the 'game-theoretical dimension'). That the two coincide is a result.

3. The notion is defined with respect to an information set.
4. The strictness requirement of dominant manipulation might apply to another than the actual profile.
5. The strictness requirement of dominant manipulation need not apply to all profiles in the information set, as in knowledge of manipulation (both *de re* and *de dicto*).
6. Strictness holds for all profiles in the information set, and therefore for some.

$$\square$$

Given Proposition 11.2, Proposition 9 stating that knowledge of manipulation is impossible in the hypercube, for majority voting, should therefore be credited to [12] who proves that dominant manipulation is impossible under common ignorance of others' preferences, for a variety of voting rules (their results were subsequently strengthened in [27]). Proposition 11.2 is somewhat surprising: it says that whether a preference is dominant does not depend on what you know of other voters' preferences. It holds given common knowledge of the profile iff it holds given common ignorance (except one's own preference) of the profile. Dominant strategy as in Def. 3 seems too strong to be useful in voting theory, as it does not even rule out that everybody except you acts against their interests.

In view of Proposition 11.6, an alternative designation for a (*de re*) known manipulation is *strongly dominant manipulation*.

5   Equilibrium and knowledge

5.1   Conditional equilibrium

Determining equilibria under incomplete knowledge comes down to decision taking under incomplete knowledge. Therefore we have to choose a decision criterion. Expected utility makes no sense here, because we didn't start with probabilities over profiles in the first place, nor with utilities. In the absence of prior probabilities, the following three criteria make sense. (i) The *insufficient reason* (or *Laplace*) criterion considers all possible states in a given situation as equiprobable. This criterion was used in [1] to determine equilibria of certain (Bayesian) games of imperfect information. (ii) The *minimax regret* criterion selects the decision minimizing the maximum utility loss, taken over all possible states, compared to the best decision, had the voter known the true state. (iii) The *pessimistic* (or *Wald*, or *maximin*) criterion compares decisions according to their worst possible consequences. The latter criterion, that we also call *risk averse*, is one that fits well our probability-free and utility-free model; this was also the criterion chosen in [12, 23]. The only assumption here is that the probability distribution is positive in all states. We now fix this

criterion for the rest of the paper. Pessimistic, optimistic, and other criteria only assuming positive probability are applied to social choice settings in [24]. Meir ([23]) also considers the minimax-regret criterion.

In the presence of knowledge, and on the assumption that voters know their own preference (so that, in game theoretical terms, the payoff function is uniform throughout an agent's information set), the definition of an equilibrium extends naturally. For each agent, the combination of an agent i and an information set $[s]_{\sim_i}$ for that agent (for some state s in the knowledge profile) defines a so-called virtual agent: we model these imperfect information games as Bayesian games [18]. Each virtual agent has the same set of strategies as the 'original' agent. An equilibrium is then a profile of strategies such that none of the virtual agents has an interest to deviate. An alternative way to present a Bayesian game, applied in [1], is to change the set of strategies instead of the set of agents. Instead of each agent in each information set (a 'virtual agent') choosing a strategy among the set of strategies, we have each agent choosing a *conditional strategy* among the larger set of conditional strategies, where conditions correspond to the information sets. We also follow that presentation for voting.

For risk-averse voters knowing their own preferences we can effectively determine if a conditional profile is an equilibrium without taking probability distributions into account, unlike in the more general setting of Bayesian games that it originates with.

For any $\mathcal{C}' \subseteq \mathcal{C}$, $\min_i \mathcal{C}'$ is the (unique) $c \in \mathcal{C}'$ such that $c' \succeq_i c$ for all $c' \in \mathcal{C}'$. Let $\mathcal{P}$ be a set of profiles, $\succ_i' \in O(\mathcal{C})$, and F a voting rule, then $\min_i F(\mathcal{P})$ denotes $\min_i \{F(\succ) \mid \succ \in \mathcal{P}\}$, and $\min_i F(\mathcal{P}_{-i}, \succ_i')$ denotes $\min_i \{F(\succ_{-i}, \succ_i') \mid \succ \in \mathcal{P}\}$.

Definition 12 (Pessimistic manipulation)   Given is a profile model $\mathcal{M} = (S, \sim, \pi)$, $s \in S$ with $\pi(s) = \succ$, and voting rule F. The worst outcome for voter i in information set $\pi([s]_{\sim_i})$ is $\min_i F(\pi([s]_{\sim_i}))$. Preference $\succ_i' \in O(\mathcal{C})$ is a *pessimistic manipulation* for voter i of $\pi([s]_{\sim_i})$ (or: of knowledge profile $\mathcal{M}_s$) iff

$$\min_i F(\pi([s]_{\sim_i})_{-i}, \succ_i') \ \succ_i \ \min_i F(\pi([s]_{\sim_i})).$$

Definition 13 (Conditional preference, conditional profile, conditional equilibrium) Given is a profile model $\mathcal{M} = (S, \sim, \pi)$ and voting rule F. For each voter i, a *conditional preference* is a function $[\succ]_i : S/\sim_i \rightarrow O(\mathcal{C})$ that assigns a preference to each information set for that voter. A *conditional profile* is a function from voters to conditional preferences. A conditional profile is a *conditional equilibrium* iff no agent has a pessimistic manipulation of any of its information sets.[2]

In the situation without uncertainty, given n voters, a profile and a voting rule determine a winner. In the strategic game matrix, the outcome is the n-tuple of values

---

[2]If all states are considered equiprobable (the Laplace criterion), a sufficient (but not necessary) condition for a conditional profile to be a conditional equilibrium is that *no agent has a **dominant** manipulation of any of its information sets*. We have not investigated this further.

(payoffs) of that winner for each voter, and to determine if it is an equilibrium we compare the value for any voter i with the value when i had voted differently: the value should not be higher. The outcome of a conditional profile is not an n-tuple of values, but an n-tuple of m-tuples or vectors $(x_i^1, \ldots, x_i^m)$, where voter i has m information sets and where $x_i^1, \ldots, x_i^m$ are expected outcomes. These vectors are unordered, so we have to compute equilibria differently. For example, given a voter 7 with two information sets x and y, we cannot say which of payoffs $(0, 1)$ and $(1, 0)$ she prefers. But we can say that virtual voter $(7, x)$ prefers the second (wherein she gets 1) over the first (wherein she gets 0), and that virtual voter $(7, y)$ prefers the first over the second. This merely is the Bayesian game calculation of equilibrium for virtual agents.

A notable fact, that we consider a main result of our contribution, is that:

Proposition 14   States with the same profile can have different conditional equilibria.

Proof   We prove this by example, in the next subsection. The reader may wish to verify in Figure 3 in that subsection: that Sunil (2) has the same preference in state u as in state v, that voting for Cars (c) is not in a conditional equilibrium in u, whereas voting for Cars is in a conditional equilibrium in v. $\square$

In other words, Proposition 14 states that even if all voters have the same preferences, then when their knowledge about others' preferences is different, their manipulative behaviour may also be different. Readers who find this result obvious may wish to skip the next subsection and proceed with Section 6 on revealing voting preferences.

5.2   Examples of conditional equilibria in plurality voting

We recall Example 7 about Leela (1) and Sunil (2) voting, by plurality, for an animated movie that may be a (Alice), b (Brave), or c (Cars), where Leela's preference is $a \succ_1 b \succ_1 c$ and Sunil's preference is $c \succ_2 b \succ_2 a$. We further assume that mother Devi, the central authority, has tie-breaking preference $b \succ_{\text{tie}} a \succ_{\text{tie}} c$. We present equilibria when there is: no uncertainty, uncertainty between two states with different profiles, and different kinds of uncertainty between three states (for two profiles).

No uncertainty   We express the payoffs for both voters by their ranking (0, 1, or 2) for the winner. As this is majority voting, preference relations with the same most preferred candidate are indistinguishable. So, 'Leela votes $a \succ_1 b \succ_1 c$' and 'Leela votes $a \succ_1 c \succ_1 b$' can both be represented by 'Leela votes a'. (Given this identification, we call a a vote and not a preference (relation).) This simplifies the outcomes matrix and the payoff matrix. If 1 votes for her preference a and 2 votes for his preference c, then the tie-breaking preferences determines a as the winner, 2's least

preferred candidate. A strategic vote of 2 for candidate b makes b win, a better outcome for voter 2. Equilibrium pairs of votes are (a, b) and (b, b). For voter 1, voting a is dominant.

1	2
a	c
b	b
c	a

$\succ$

1	2
c	c
b	b
a	a

$\succ'$

$1\backslash 2$	a	b	c
a	a	b	a
b	b	b	b
c	a	b	c

$1\backslash 2$	a	b	c
a	2.0	1.1	2.0
b	1.1	1.1	1.1
c	2.0	1.1	0.2

$1\backslash 2$	a	b	c
a	0.0	1.1	0.0
b	1.1	1.1	1.1
c	0.0	1.1	2.2

Figure 1: Equilibria for $\succ$ and for $\succ'$. From left to right and from above to below: profile $\succ$, profile $\succ'$, the matrix with winners, the outcome matrix for profile $\succ$, and the outcome matrix for profile $\succ'$. Equilibria are boxed. We write $i.j$ instead of (i, j) to denote the values for voters 1 and 2 of the outcome of the election.

The other profile used in the examples in this section is where 1 shares the preferences of 2. This is the profile $\succ'$. Although (c, c) is an equilibrium vote for $\succ'$, there are various suboptimal equilibria. There is no dominant vote. An overview of the equilibria for $\succ$ and for $\succ'$ is in Figure 1.

Uncertainty between two profiles   Now consider the profile model consisting of two states t with profile $\succ$ and u with profile $\succ'$ and that are the same for voter 2, but different for voter 1 (the accessibility relation for voter 1 is the identity on the model and for voter 2 it is the universal relation). Figure 2 depicts that profile model, the strategic game matrix with conditional preferences and winners, and the strategic game matrix with payoffs. Conditional profiles are pairs (ij, k) where i is 1's vote in t and j is 1's vote in u, and k is 2's vote in $\{t, u\}$.

As voter 1 has two information sets, the conditional preference for 1 has two conditions, for each of which a choice between the three candidates (co-)determines the outcome of the majority vote. There are therefore 9 conditional preferences for voter 1. The matrix shows the conditional preferences for 1 by the candidate she votes for in t, followed by the candidate she votes for in u. Conditional preference xy for 1 means that in state t 1 votes x and in state u 1 votes y. The payoff matrix next to the winners matrix contains triples $ij.k$ for, in this order: the value of the worst outcome for 1 given t ($\succ$) of that conditional profile, the value of the outcome for 1 given u ($\succ'$), and the value of the worst outcome for 2 given t, u, his only information set.

For example, for conditional profile (ba, c) we get ba as the entry in the winners

matrix and (10.0) as the entry in the payoff matrix: if 1 votes b and 2 votes c then the tie ($b \succ_{\text{tie}} a \succ_{\text{tie}} c$) makes b win, value 1 for voter 1 and value 1 for voter 2 in $\succ$; if 1 votes a and 2 votes c then a wins, value 0 for voter 1 and for voter 2 in $\succ'$; the worst of 0 and 1 is 0, so the value for voter 2 of this conditional profile is 0.

1	2
a	c
b	b
c	a

$t, \succ$

—2—

1	2
c	c
b	b
a	a

$u, \succ'$

$1\backslash 2$	a	b	c
aa	aa	bb	aa
ab	ab	bb	ab
ac	aa	bb	ac
ba	ba	bb	ba
bb	bb	bb	bb
bc	ba	bb	bc
ca	aa	bb	ca
cb	ab	bb	cb
cc	aa	bb	cc

$1\backslash 2$	a	b	c
aa	20.0	11.1	20.0
ab	21.0	11.1	21.0
ac	20.0	11.1	21.0
ba	10.0	11.1	10.0
bb	11.1	11.1	11.1
bc	10.0	11.1	12.1
ca	20.0	11.1	00.0
cb	21.0	11.1	01.1
cc	20.0	11.1	02.2

Figure 2: Voter 2 is uncertain whether voter 1 prefers a over c or c over a

The equilibria are, maybe, as expected. (We only consider pure strategies.) If the profile is $\succ$ then it is still dominant for voter 1 to vote a (if the profile is $\succ'$, voting for c is not dominant for voter 1). Because voter 2 is risk averse, (c, c) is no longer an equilibrium vote in $\succ'$. As 2 is uncertain whether 1 prefers c over a or a over c, the safer (risk avoiding) strategy for 2 is now to vote b, even when 1 and 2 both prefer c. Voter 1 knows this as well.

Voter 2 does not have a dominant preference, because if he assumes that voter 1 always votes c, the best response is also to vote c and not to vote b. So this is the only case where voting b is not an equilibrium vote for 2.

Uncertainty between three states We now add further uncertainty to the two-state profile model where 2 is uncertain between profiles $\succ$ and $\succ'$. Figure 3 displays two different ways to do this. In both depicted profile models voter 1 always knows voter 2's preferences. We will show that it is not rational for 2 to behave (vote) differently in s and in t, in the first, but that it is rational for 2 to behave differently in u and in v, in the second.

Figure 3 also gives an overview of the conditional equilibria for both profile models, including the matrices with winners in order to calculate the payoffs. As it may be confusing to see three winners but four payoff values let us explain once more the mechanics of conditional profiles and conditional equilibria. For example, take the t, u, v model, with conditional profile (ac, bc) that is an equilibrium, where in the winners matrix we find bbc for that entry, and where in the payoff matrix we find 11.12. Conditional profile (ac, bc) is the conditional profile such that
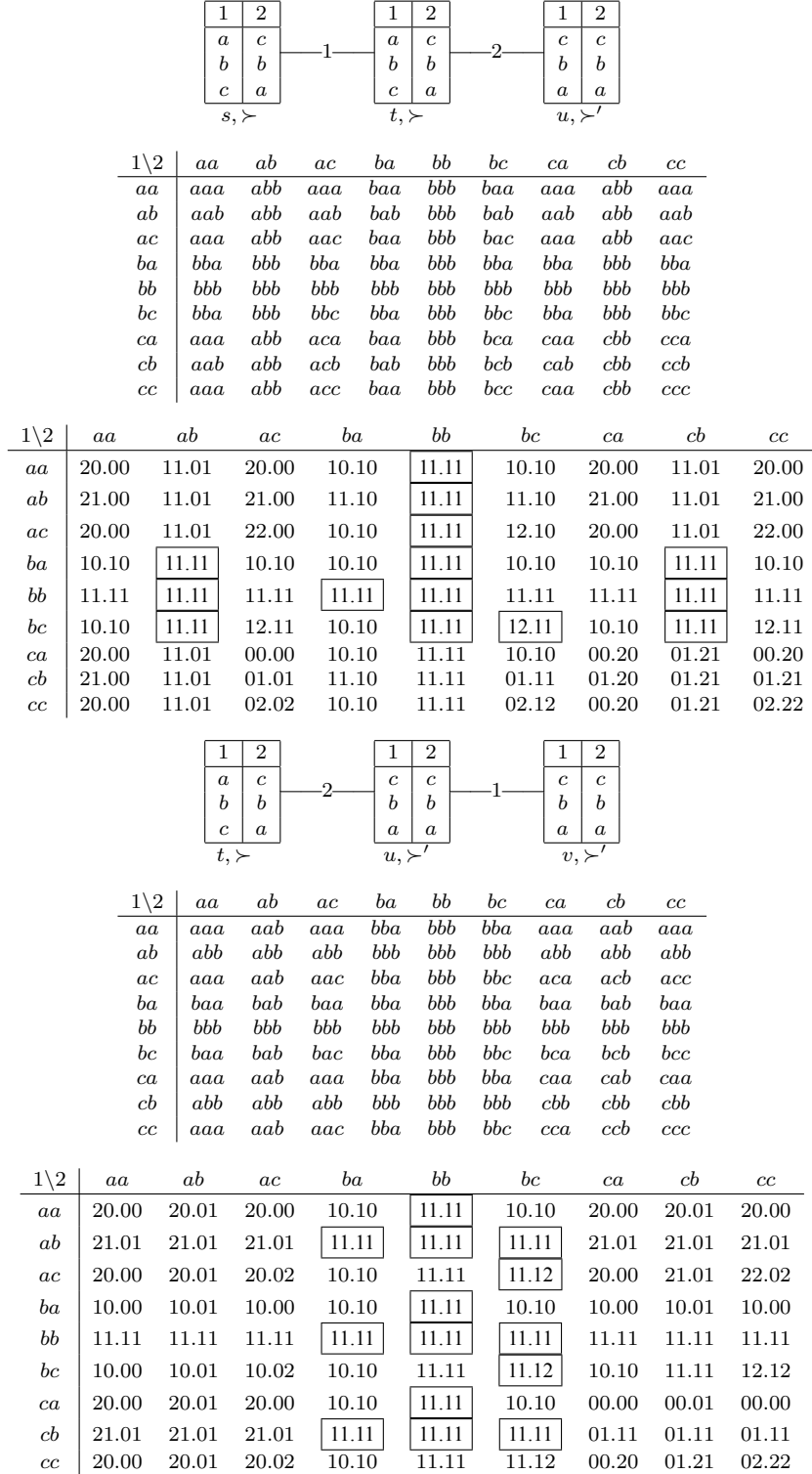
	1	2
	a	c
	b	b
	c	a

$s, \succ$ —1— $t, \succ$ —2— $u, \succ'$

$t,\succ$	1	2		$u,\succ'$	1	2
	a	c			c	c
	b	b			b	b
	c	a			a	a

1\2	aa	ab	ac	ba	bb	bc	ca	cb	cc
aa	aaa	abb	aaa	baa	bbb	baa	aaa	abb	aaa
ab	aab	abb	aab	bab	bbb	bab	aab	abb	aab
ac	aaa	abb	aac	baa	bbb	bac	aaa	abb	aac
ba	bba	bbb	bba	bba	bbb	bba	bba	bbb	bba
bb	bbb	bbb	bbb	bbb	bbb	bbb	bbb	bbb	bbb
bc	bba	bbb	bbc	bba	bbb	bbc	bba	bbb	bbc
ca	aaa	abb	aca	baa	bbb	bca	caa	cbb	cca
cb	aab	abb	acb	bab	bbb	bcb	cab	cbb	ccb
cc	aaa	abb	acc	baa	bbb	bcc	caa	cbb	ccc

1\2	aa	ab	ac	ba	bb	bc	ca	cb	cc
aa	20.00	11.01	20.00	10.10	[11.11]	10.10	20.00	11.01	20.00
ab	21.00	11.01	21.00	11.10	[11.11]	11.10	21.00	11.01	21.00
ac	20.00	11.01	22.00	10.10	[11.11]	12.10	20.00	11.01	22.00
ba	10.10	[11.11]	10.10	10.10	[11.11]	10.10	10.10	[11.11]	10.10
bb	11.11	[11.11]	11.11	[11.11]	[11.11]	11.11	11.11	[11.11]	11.11
bc	10.10	[11.11]	12.11	10.10	[11.11]	[12.11]	10.10	[11.11]	12.11
ca	20.00	11.01	00.00	10.10	11.11	10.10	00.20	01.21	00.20
cb	21.00	11.01	01.01	11.10	11.11	01.11	01.20	01.21	01.21
cc	20.00	11.01	02.02	10.10	11.11	02.12	00.20	01.21	02.22

$t,\succ$	1	2		$u,\succ'$	1	2		$v,\succ'$	1	2
	a	c			c	c			c	c
	b	b			b	b			b	b
	c	a			a	a			a	a

$t, \succ$ —2— $u, \succ'$ —1— $v, \succ'$

1\2	aa	ab	ac	ba	bb	bc	ca	cb	cc
aa	aaa	aab	aaa	bba	bbb	bba	aaa	aab	aaa
ab	abb	abb	abb	bbb	bbb	bbb	abb	abb	abb
ac	aaa	aab	aac	bba	bbb	bbc	aca	acb	acc
ba	baa	bab	baa	bba	bbb	bba	baa	bab	baa
bb	bbb	bbb	bbb	bbb	bbb	bbb	bbb	bbb	bbb
bc	baa	bab	bac	bba	bbb	bbc	bca	bcb	bcc
ca	aaa	aab	aaa	bba	bbb	bba	caa	cab	caa
cb	abb	abb	abb	bbb	bbb	bbb	cbb	cbb	cbb
cc	aaa	aab	aac	bba	bbb	bbc	cca	ccb	ccc

1\2	aa	ab	ac	ba	bb	bc	ca	cb	cc
aa	20.00	20.01	20.00	10.10	[11.11]	10.10	20.00	20.01	20.00
ab	21.01	21.01	21.01	[11.11]	[11.11]	[11.11]	21.01	21.01	21.01
ac	20.00	20.01	20.02	10.10	11.11	[11.12]	20.00	21.01	22.02
ba	10.00	10.01	10.00	10.10	[11.11]	10.10	10.00	10.01	10.00
bb	11.11	11.11	11.11	[11.11]	[11.11]	[11.11]	11.11	11.11	11.11
bc	10.00	10.01	10.02	10.10	11.11	[11.12]	10.10	11.11	12.12
ca	20.00	20.01	20.00	10.10	[11.11]	10.10	00.00	00.01	00.00
cb	21.01	21.01	21.01	[11.11]	[11.11]	[11.11]	01.11	01.11	01.11
cc	20.00	20.01	20.02	10.10	11.11	11.12	00.20	01.21	02.22

Figure 3: Conditional equilibria for profile models where two states have the same profile

- If 1 prefers a (i.e., in state t) then she votes a, and if 1 prefers c (i.e., in states u, v) then she votes c.
- If 2 is uncertain whether 1 prefers a (i.e., in states t, u) then he votes b, and if 2 knows that 1 prefers c (i.e., in state v) then he votes c.

The winners in states t, u, v of these conditional preferences are, respectively, b, b, c. If the state is t, then 1 votes a and 2 votes b, so b wins. If u, then 1 votes c and 2 votes b, so b wins. If v, then 1 votes c and 2 votes c, so c wins.

The payoff entry is 11.12 because: for voter 1 in state t, the worst (and only) outcome is b with value 1, for voter 1 in states u, v the worst of b and c, with values 1 and 2, is (also) 1; for voter 2 in states t, u the worst of b and b, both with value 1, is 1, and for voter 2 the worst and only outcome in state v is c with value 2.

Conditional profile (ac, bc) is an equilibrium. For this, we have to check four virtual voters. For example, voter 1 in state t cannot do better, because the first digit of the payoff entries for profiles (bc, bc), (cc, bc) is not greater than 1; voter 1 in class $\{u, v\}$ cannot do better: we check the second digit of the entries for conditional profiles (aa, bc) and (ab, bc). Voter 2 cannot do better in t, u, check the third digit in entries for (ac, ac) and (ac, cc); and voter 2 also cannot do better in v, in which case we check the fourth digit in the entries for (ac, ba) and (ac, bb). We are done.

In the s, t, u case, it does not make a difference to voter 2 whether he knows voter 1's preferences. If the profile is $\succ$, voter 1 *knows* that voting for a is dominant. On that assumption, voter 2 should vote b, such that b wins. Indeed, in almost all equilibria (except (bc, bc)), b wins and the payoff is 1 for both voters. Unlike for the two-state example, where in all equilibria voter 2 votes b, there are now equilibria wherein voter 2 does not vote b. However, these are not really interesting, as 1 votes b in these, which is dominated by 1 voting a.

On the other hand, in the t, u, v case, it makes a difference to voter 2 if 1 is uncertain or not. There are equilibria wherein both players vote c in state v, namely (ac, bc) and (bc, bc). Whereas there is no equilibrium wherein both players vote c in state u, even though that would have been just as much in their interest. We can easily justify this result by our intuitions. If voter 2 is uncertain about 1's preferences, the worst-case avoiding strategy remains voting b. If voter 2 knows that 1's preferences are his own, even 1's uncertainty is not enough to make him change his vote. The same cannot be said for voter 1 in that state v, she has to weigh the odds against voter 2 playing safe and voting b instead; but either way, voting c then also gives her best result. So voter 1 should be indifferent between b and c when arguing from that worst-case scenario, and this is indeed the case: (ab, bc) and (bb, bc) are also equilibria.

6   Revealing voting preferences

We can extend the setting for the interaction of voting preferences and knowledge of the previous sections with operations wherein voters are informed of other's

preferences, thus reducing ('updating') their uncertainty. An obvious choice for such updates is the *public announcement* [26] of propositions about profiles. A public announcement can be modelled as an operation $\mathcal{M}_s \mapsto (\mathcal{M}|T)_s$, where $T \subseteq S$ is the denotation in $\mathcal{M}$ of a proposition about profiles φ, and $\mathcal{M}|T$ means model restriction to subdomain T. In that case we also write $(\mathcal{M}|\varphi)_s$.

Given a knowledge profile $\mathcal{M}_s$, the precondition for execution of the operation *public announcement of φ* (or *update with φ*) is that φ is true in $\mathcal{M}_s$, and the way to execute it is to restrict the model $\mathcal{M}$ to all the states where φ is true. We can then investigate the truth of propositions about profiles in that model restriction. This allows us to evaluate propositions about profiles of shape 'after update with φ, ψ (is true),' such as: 'After voter 1 reveals her preference, voter 2 knows that he has a manipulation'.

All this is embodied in the following definition (see Section 7 for a formal version).

Definition 15 (Updated knowledge profile)   Let $\mathcal{M}_s$ be a knowledge profile, where $\mathcal{M} = (S, \sim, \pi)$, and let φ be a proposition about profiles with denotation $S' \subseteq S$ and such that $s \in S'$. Then the *updated knowledge profile* $(\mathcal{M}|\varphi)_s$ is defined as $\mathcal{M}|\varphi = (S', \sim', \pi')$ where $\sim'_i = \sim_i \cap (S' \times S')$ and for all $\succ \in O(\mathcal{C})^{\mathcal{N}}$, $\pi'(\succ) = \pi(\succ) \cap S'$.

Let ψ also be a proposition about profiles. In $\mathcal{M}_s$ it is true that ψ after update with φ, iff whenever φ is true in $\mathcal{M}_s$, ψ is true in $(\mathcal{M}|\varphi)_s$.

A public announcement is considered information coming from an outside source (for example, a central authority) and that is reliable. However, it is common in dynamic epistemic logic to model a public announcement φ by an agent a (an inside source, so to speak, that is modelled in the system) as the public announcement of $K_a\varphi$. This makes it possible to formalize that a voter reveals her preference to the other voters, as above. But given this identification, a voter revealing her preferences to me is indistinguishable in the logical analysis from someone else revealing to me that voter's preferences. Yet another situation is when a voter reveals her preferences *only* to another voter but not to all voters. This is not a public announcement but a *private* announcement. Formalizing this is quite doable but requires a more complex analysis. Finally, announcements can be *insincere*, or *lies*. This also requires a more complex dynamic epistemic analysis.

We proceed with some results. Clearly, manipulations and equilibrium profiles are preserved after update, as these only depend on the profile of the actual state, that is preserved after any (truthful) update. These are not epistemic notions. For those, we have that:

Proposition 16   Knowledge of manipulation is preserved after update.

Proof   In any state of the information set of the voter knowing the manipulation, the profile of that state has a manipulation, by Def. 8. This is a universal property that is

preserved after update. This holds for *de re* as well as *de dicto* knowledge. □

Proposition 17 Dominant manipulation is not preserved after update.

Proof This is an existential property that may not be preserved, namely if the (existential) strictness requirement is only satisfied in states that are removed in the update. □

To investigate how conditional equilibria evolve after updates we first define the update of a conditional profile.

Definition 18 (Updated conditional profile) Let profile model $\mathcal{M} = (S, \sim, \pi)$ and conditional profile $[\succ]$ be given, where $[\succ]_i : S\backslash\sim_i \to O(\mathcal{C})$ are conditional preferences. Let φ be a proposition about profiles such that $\mathcal{M}|\varphi = (S', \sim', \pi')$. Then the *updated conditional profile* $[\succ']$ (w.r.t. $\mathcal{M}|\varphi$) consists of conditional preferences $[\succ']_i : S'\backslash\sim'_i \to O(\mathcal{C})$ defined as: for all $s \in S'$, $[\succ']_i([s]_{\sim'_i}) = [\succ]_i([s]_{\sim_i})$.

A preference $[\succ]_i : S\backslash\sim_i \to O(\mathcal{C})$ may be affected in three different ways in an update:

- An information set for voter i disappears, namely if none of its states satisfies the update φ. The updated preference then has one less condition (the virtual voter $(i, [s]_{\sim_i})$ ceases to exist).
- An information set for voter i shrinks, because some states satisfy φ and others do not satisfy φ. We then have that $[s]_{\sim'_i} \subset [s]_{\sim_i}$. This may affect the value for i of preference $\succ_i$: states with minimal value may have been removed, namely if $\min_i F(\pi([s]_{\sim_i})) \prec_i \min_i F(\pi'([s]_{\sim'_i}))$.
- An information set for i remains the same, because all of its states satisfy φ. The expected worst outcome for i remains the same.

Proposition 19 Conditional equilibrium is not preserved after update.

Proof An information update may affect an equilibrium as follows. The outcome for a virtual voter $(i, [s]_{\sim_i})$ casting vote $\succ_i$ in the equilibrium is the worst outcome (winner) for voter i in information set $[s]_{\sim_i}$. This payoff value of that winner is at least as good as the worst outcome in information set $[s]_{\sim_i}$ for any other preference $\succ''_i$. This is fragile and not preserved after update. More precisely, we may have that (we recall that for all $\succ' \in \pi([s]_{\sim_i}), \succ'_i = \succ_i$):

$$\min_i F(\pi([s]_{\sim_i})) \quad \succeq_i \quad \min_i F(\pi([s]_{\sim_i})_{-i}, \succ''_i)$$
$$\text{whereas} \quad \min_i F(\pi'([s]_{\sim'_i})) \quad \prec_i \quad \min_i F(\pi'([s]_{\sim'_i})_{-i}, \succ''_i)$$

where as before π' and $\sim'_i$ are the valuation and accessibility relation in the updated profile model; and this requires a $\succ' \in \pi([s]_{\sim_i}\backslash[s]_{\sim'_i})$ such that:

$$F(\succ') \quad \succ_i \quad F(\succ'_{-i}, \succ''_i) \qquad \Box$$

Proposition 20 (Not being a conditional equilibrium) is not preserved after update.

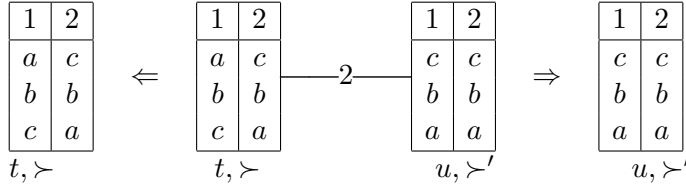**Proof** This requires a preference $\succ_i''$ for which
$$\min_i F(\pi([s]_{\sim_i}) \prec_i \min_i F(\pi([s]_{\sim_i})_{-i}, \succ_i'')$$
$$\text{and } \min_i F(\pi'([s]_{\sim_i'})) \succeq_i \min_i F(\pi'([s]_{\sim_i'})_{-i}, \succ_i'').$$
$\square$

So conditional equilibria can both disappear and appear after updates (see also Example 21 below). It is unclear to us if there are general patterns here. However, additional strategic behaviour comes into the picture with these negative results. An update may consist of a voter revealing her voting preferences. It may be that this voter's sincere preference is not part of an equilibrium conditional profile, but that after this voter reveals her sincere preference, the updated conditional profile is an equilibrium. This interaction between *strategic voting* and *strategic communication* may be of interest.

Example 21 We recall the two-state profile model presented in Section 5.2 and displayed again below. We now add dynamics to this example: voter 1 informs voter 2 of her true preference. This is also displayed below.

$$
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
a & c \\
b & b \\
c & a \\
\hline
\end{array}
\quad t, \succ
\qquad \Longleftarrow \qquad
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
a & c \\
b & b \\
c & a \\
\hline
\end{array}
\quad t, \succ
\quad\text{---}2\text{---}\quad
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
c & c \\
b & b \\
a & a \\
\hline
\end{array}
\quad u, \succ'
\qquad \Longrightarrow \qquad
\begin{array}{|c|c|}
\hline
1 & 2 \\
\hline
c & c \\
b & b \\
a & a \\
\hline
\end{array}
\quad u, \succ'
$$

In state t, after voter 1 informs voter 2 of her true preference $a \succ_1 c$, no uncertainty remains, and 1 and 2 commonly know that the profile is $\succ$. From Section 5.2 we further recall that equilibrium votes for $\succ$ are $(a, b), (b, b)$, and that conditional equilibria for the t, u profile model have shape (xy, b), where x is voter 1's preference in t and y is voter 1's preference in u, and where all but cc are equilibria. (See the payoff matrix in Figure 2.)

We can now observe that all conditional equilibria are preserved after update. For example, given conditional profile (bc, b), the updated conditional profile according to Def. 15 is (b, b). There is therefore no strategic incentive for voter 1 to inform voter 2 in state t.

On the other hand, in state u voter 1 has an incentive to make her preference $c \succ_1 a$ known to 2. In the model with $\succ$ and $\succ'$, there is no equilibrium wherein 2 votes c. But after 1 informs 2 that her preferences are the same as his preferences, (c, c) is an equilibrium. Most conditional equilibria of the two-state model are preserved, but not those where 1 votes c. For example, (ac, b) was an equilibrium, but the updated profile (c, b) is not an equilibrium.

Voter 1 has a strategic interest to reveal her preferences to voter 2 in state u because subsequently she expects the outcome of the vote to be better. Before the

update she expected b to win, after update she expects c to win, and $c \succ_1 b$. This demonstrates that:

> *When there is uncertainty about votes, voters have different ways of acting strategically: voting strategically or strategically revealing voting preferences.*

7   A logic of knowledge and voting

Given the set $\mathcal{N}$ of agents, the set of profiles $O(\mathcal{C})^{\mathcal{N}}$, that serve as *propositional variables*, and a voting rule F, we now define a logical language and semantics.

Definition 22 (Logical language)   The language $\mathcal{L}$ is defined as

$$\varphi ::= \ \succ \ | \ \neg\varphi \ | \ \varphi \wedge \varphi \ | \ K_i\varphi \ | \ [\varphi]\varphi$$

where $i \in \mathcal{N}$ and $\succ \ \in O(\mathcal{C})^{\mathcal{N}}$.

An element of the language is a *formula*, and φ is a *formula variable*; $K_i\varphi$ stands for 'voter i knows that φ'; $[\varphi]\psi$ stands for 'after (public) announcement of φ, ψ (is true)'.

Definition 23 (Semantics)   Let $\mathcal{M}_s$ be a knowledge profile, where $\mathcal{M} = (S, \sim, \pi)$. The interpretation of formulas in a knowledge profile is defined as follows:

$$
\begin{aligned}
\mathcal{M}_s &\models \ \succ & \text{iff} \quad & \pi(s) = \ \succ \\
\mathcal{M}_s &\models \neg\varphi & \text{iff} \quad & \mathcal{M}_s \not\models \varphi \\
\mathcal{M}_s &\models \varphi \wedge \psi & \text{iff} \quad & \mathcal{M}_s \models \varphi \text{ and } \mathcal{M}_s \models \psi \\
\mathcal{M}_s &\models K_i\varphi & \text{iff} \quad & \text{for every } t \text{ such that } s \sim_i t, \mathcal{M}_t \models \varphi \\
\mathcal{M}_s &\models [\varphi]\psi & \text{iff} \quad & \mathcal{M}_s \models \varphi \text{ implies } (\mathcal{M}|\varphi)_s \models \psi
\end{aligned}
$$

where $\mathcal{M}_s \not\models \varphi$ stands for 'not ($\mathcal{M}_s \models \varphi$)', and where $\mathcal{M}|\varphi = (S', \sim', \pi')$ such that $S' = \{t \in S \mid \mathcal{M}_t \models \varphi\}$, $\sim'_i = \ \sim_i \cap (S' \times S')$, and for all $s \in S'$, $\pi'(s) = \pi(s)$. If $\mathcal{M}_s \models \varphi$ for all $s \in S$, we write $\mathcal{M} \models \varphi$ (φ is *valid on $\mathcal{M}$*) and if this is the case for all $\mathcal{M}$, we say that φ is *valid*, and we write $\models \varphi$. *The logic of knowledge and voting* is the set of validities for the class of profile models.

Although the voting function F is not used in the syntax or in the semantics, it will later be used in the logical abbreviations introduced to formalize concepts such as conditional equilibrium.

Profile models are standard Kripke models, but with valuations that satisfy special conditions. We now present principles that are valid on the class of profile models, and that will feature as axioms in the proof system. Let $\overline{\bigvee}$ denote exclusive disjunction.

$$
\begin{aligned}
\mathbf{P} \ &: \ \overline{\bigvee} O(\mathcal{C})^{\mathcal{N}} \\
\mathbf{N} \ &: \ \bigwedge_{i \in \mathcal{N}} \bigwedge_{\succ_i \in O(\mathcal{C})} (\succ_i \rightarrow K_i \succ_i)
\end{aligned}
$$

Axiom **P** spells out that only one profile is true in each state. Axiom **N** says that voters know their own preferences.

Proposition 24   The axioms **P** and **N** are valid on the class of profile models.

Proof   Axiom **P** is valid as the valuation π is a function from states to profiles. Axiom **N** formalizes the constraint on knowledge profiles that $s \sim_i t$ implies $\pi(s)_i = \pi(t)_i$. $\hspace{1cm}\square$

Proposition 25   The logic of knowledge and voting has a complete axiomatization.

Proof   As the axiomatization of the logic of knowledge and voting we propose the axiomatization of public announcement logic, that is standard and that we assume to be known [26], and to which we add the axioms **P** and **N**. This axiomatization is with respect to the class of profile models (see the definition of the logic, above). For the soundness we refer to the soundness of public announcement logic and the validity of **P** and **N** (Prop. 24). For the completeness we observe that the canonical model to determine the completeness of the logic without public announcements is a profile model (Prop. 24, again), and that the completeness of the logic with announcements is as usual (see [13]) obtained because every formula is equivalent to one without announcements (the axioms are rewriting rules, pushing all logical connectives beyond announcements, such as $[\varphi](\psi \wedge \chi) \leftrightarrow ([\varphi]\psi \wedge [\varphi]\chi)$). $\hspace{1cm}\square$

We proceed to formalize the concepts involving knowledge and voting. The main definition of interest is that of conditional equilibrium. We start with some abbreviations.[3]

$$
\begin{aligned}
\succ_i & := & \bigvee \{\succ' \in O(\mathcal{C})^{\mathcal{N}} \mid \succ_i = \succ'_i\} \\
a \succ_i b & := & \bigvee \{\succ' \in O(\mathcal{C})^{\mathcal{N}} \mid a \succ'_i b\} \\
a & := & \bigvee \{\succ \in O(\mathcal{C})^{\mathcal{N}} \mid F(\succ) = a\} \\
(\succ_{-i}, \succ'_i) & := & \bigvee \{\succ'' \in O(\mathcal{C})^{\mathcal{N}} \mid (\succ_{-i}, \succ'_i) = \succ''\} \\
F(\succ') \succ_i F(\succ'') & := & ((\succ' \to a) \wedge (\succ'' \to b)) \to (a \succ_i b)
\end{aligned}
$$

We emphasize the role of the background parameter F in these abbreviations. Firstly, the 'formula' a above (third item) stands for 'candidate a wins the election'. The definiens uses F. Formula a is therefore true in a state s with profile $\pi(s)$ such that $F(\pi(s)) = a$. Secondly, the use of F in the definiendum of the final abbreviation (fifth item), $F(\succ') \succ_i F(\succ'')$, is therefore proper, because the definiens refers to these winners, namely in $\succ' \to a$ and $\succ'' \to b$: that a and b are abbreviations of formulas that use F.

Example 26   We demonstrate the logic reusing Example 7 about Leela and Sunil.

---

[3]In the 4th abbreviation, note that the $\succ''$ such that $(\succ_{-i}, \succ'_i) = \succ''$ is unique, and that $\bigvee \{\succ''\} = \succ''$.

- Leela prefers Alice over Cars: $\mathcal{M}_t \models a \succ_1 c$
- Sunil does not know that Leela prefers Alice over Cars: $\mathcal{M}_t \models \neg K_2(a \succ_1 c)$
- Leela knows Sunil's preference, but Leela is uncertain whether Sunil knows her preference: $\mathcal{M}_t \models K_1 \succ_2 \wedge \neg(K_1 K_2 \succ_1 \vee K_1 \neg K_2 \succ_1)$
- Sunil does not know that Leela prefers Alice over Cars, but after he was told so, he knows it: $\mathcal{M}_t \models \neg K_2(a \succ_1 c) \wedge [a \succ_1 c]K_2(a \succ_1 c)$

Using the above abbreviations and trivial ones like $a \succeq_i b := \neg(b \succ_i a)$ we now have that:

- Voter i has a manipulation of profile $\succ$:
  $\bigvee_{\succ'_i \in O(\mathcal{C})}(F(\succ_{-i}, \succ'_i) \succ_i F(\succ))$
- Voter i has a manipulation $\succ'_i$ in profile $\succ$:
  $F(\succ_{-i}, \succ'_i) \succ_i F(\succ)$
- Voter i has a dominant manipulation $\succ'_i$ in profile $\succ$:
  $\bigwedge_{\succ'' \in O(\mathcal{C})^{\mathcal{N}}}(F(\succ''_{-i}, \succ'_i) \succeq_i F(\succ'')) \wedge \bigvee_{\succ'' \in O(\mathcal{C})^{\mathcal{N}}}(F(\succ''_{-i}, \succ'_i) \succ_i F(\succ''))$
- Voter i with preference $\succ_i$ knows de dicto that she has a manipulation:
  $K_i \bigvee_{\succ'_i \in O(\mathcal{C})}(\succ'' \to (F(\succ''_{-i}, \succ'_i) \succ_i F(\succ'')))$
- Voter i with preference $\succ_i$ knows de re that she has a manipulation:
  $\bigvee_{\succ'_i \in O(\mathcal{C})} K_i(\succ'' \to (F(\succ''_{-i}, \succ'_i) \succ_i F(\succ'')))$
- Profile $\succ$ is an equilibrium profile:
  $\bigwedge_{i \in \mathcal{N}} \bigwedge_{\succ'_i \in O(\mathcal{C})}(F(\succ) \succeq_i F(\succ_{-i}, \succ'_i))$
- A conditional equilibrium (where $\mathcal{J}$ is defined below):

$$\bigwedge_{\mathcal{J}}\left((\bigwedge_{i \in \mathcal{N}} \varphi_i^{j(i)}) \to \bigwedge_{k \in \mathcal{N}} \bigwedge_{\succ'_k \in O(\mathcal{C})}\left(F(\bigwedge_{i \in \mathcal{N}} \succ_i^{j(i)}) \succeq_k^{j(i)} F((\bigwedge_{i \in \mathcal{N}} \succ_i^{j(i)})_{-k}, \succ'_k)\right)\right)$$

In the formalization of knowledge *de re* and *de dicto*, condition '$\succ'' \to$' guarantees that only profiles $\succ''$ in the information set $\pi([t]_{\sim_i})$ of a knowledge profile $\mathcal{M}_t$ are selected.

The index set $\mathcal{J}$ in the definition of conditional equilibrium ranges over *distinguishing formulas* $\varphi_i^{j(i)}$ for all (n) voters and for all information sets of those voters, i.e., $\mathcal{J} := \{(j(1), \ldots, j(n)) \mid 1 \leq j(1) \leq \max(1), \ldots, 1 \leq j(n) \leq \max(n)\}$ where each voter i has $\max(i)$ information sets; and where for condition $\varphi_i^{j(i)}$ voter i has preference $\succ_i^{j(i)}$. A distinguishing formula is only true in that information set of the voter and else false. As the domain is finite, such distinguishing formulas exist. The distinguishing formulas *cover* the profile model, i.e., let the information sets for a voter in a given profile model be numbered $1, \ldots, \max$, then there are formulas $\varphi^1, \ldots, \varphi^{\max}$ such that $\bigvee_{j=1 \ldots \max} \varphi^j$ and $\bigwedge_{j=1 \ldots \max}(\varphi^j \to \neg \varphi^{j+1})$ (where $\max + 1 = 1$) are valid on the profile model. The formalization of conditional equilibrium then simply spells out the equilibrium for the Bayesian game with virtual agents ('virtual voters') $(i, [s]_{\sim_i})$ instead of voters i. Note that $\bigwedge_{i \in \mathcal{N}} \succ_i^{j(i)}$ represents a profile: a validity of the logic is that $\succ \leftrightarrow \bigwedge_{i \in \mathcal{N}} \succ_i$.

8   Conclusion and further research

We presented a logic for the interaction of voting and knowledge. The semantic primitive is the knowledge profile: a profile including uncertainty of voters about the profile. We defined de re knowledge of manipulation and de dicto knowledge of manipulation, and the notion of conditional equilibrium for risk-averse voters. We modelled the dynamics of knowledge, such as voters revealing preferences, and its effects on knowledge of manipulation and on conditional equilibrium, where we proved that knowledge of manipulation is preserved after such updates whereas a conditional equilibrium may not be preserved. Finally, we formalized the system in a dynamic epistemic logic.

Our goal was to present a *minimally* meaningful extension of works on uncertainty in voting, namely that permits formalizing higher-order uncertainty and updates of uncertainty. Many further extensions are conceivable.

Additionally to uncertainty over the preferences of other voters one can consider uncertainty over the voting function. Instead of profile models consisting of states with a profile for each such state, we would then need 'voting models' consisting of states with a profile and a voting function for each state. Notions like pessimistic manipulation, conditional preference, and conditional equilibrium generalize straightforwardly to this setting. The logic would need an additional axiom to describe that only one voting function can be associated with a given state.

Apart from voters $1, \ldots, n$ it is convenient to distinguish a designated additional agent 0 who is the *central authority*, or *chair*. The universal relation on a knowledge profile model can then be seen as the indistinguishability relation of that agent 0. This opens the door to the logical modelling of well-studied problems in computational social choice, such as control by the chair, or determining possible winners.

Our results are for any amount of voters but all our examples were for at most three voters. This was for the purpose of presentation. Still, in realistic settings the power of individual voters is very limited. However, our results seem also to meaningfully apply to (few) *coalitions* of voters. In voting theory, the power of a coalition means the power of a set of agents that can decide on a joint action as a result of communication between them. Under conditions of uncertainty about profiles this is therefore a function of the *distributed knowledge* of that coalition. In a knowledge profile, the indistinguishability relation for a coalition is the intersection of the indistinguishability relations for the individuals in the coalition.

We modelled *knowledge* of preferences, not *belief*. Unlike knowledge, beliefs may be incorrect. Somewhat similarly, unlike truthful announcements, insincere announcements (e.g., lying about your true preferences in a voting poll) may lead to false beliefs. Conceptually, the interaction between belief and voting is much more complex than between knowledge and voting. Technically, there may be fewer issues, for example, the same logic as in Section 7 can be used with minor adjustments.

References

[1]   T. Ågotnes and H. van Ditmarsch, 2011, "What will they say? — Public announcement games", *Synthese*, **179(S.1)**: 57–85.

[2]   T. Ågotnes, W. van der Hoek and M. Wooldridge, 2011, "On the logic of preference and judgment aggregation", *Autonomous Agents and Multi-Agent Systems*, **22(1)**: 4–30.

[3]   T. Ågotnes, W. van der Hoek and M. Wooldridge, 2006, "Towards a logic of social welfare", *Proc. of 7th LOFT*, pp. 1–10.

[4]   Z. Bakhtiari, 2017, *The Dynamics of Incomplete and Inconsistent Information: Applications of logic, algebra and coalgebra*, phdthesis, Université de Lorraine, https://tel.archives-ouvertes.fr/tel-01680340.

[5]   S. Barberà, A. Bogomolnaia and H. van der Stel, 1998, "Strategy-proof probabilistic rules for expected utility maximizers", *Mathematical Social Sciences*, **35(2)**: 89–103.

[6]   J. Bartholdi, C. Tovey and M. Trick, 1989, "The computational difficulty of manipulating an election", *Social Choice and Welfare*, **6(3)**: 227–241.

[7]   J. van Benthem, 2001, "Games in dynamic epistemic logic", *Bulletin of Economic Research*, **53(4)**: 219–248.

[8]   D. Bergemann and S. Morris, 2005, "Robust mechanism design", *Econometrica*, **73(6)**: 1771–1813.

[9]   C. Boutilier and J. Rosenschein, 2016, "Incomplete information and communication in voting", in F. Brandt et al. (eds.), *Handbook of Computational Social Choice*, Cambridge: Cambridge University Press.

[10]   S. Chopra, E. Pacuit and R. Parikh, 2004, "Knowledge-theoretic properties of strategic voting", *Proc. of 9th JELIA*, pp. 18–30.

[11]   G. Ciná and U. Endriss, 2016, "Proving classical theorems of social choice theory in modal logic", *Autonomous Agents and Multi-Agent Systems*, **30(5)**: 963–989.

---

[4]The following results originate in the TARK version: Proposition 3 of TARK is essentially the same as our Proposition 17, and Proposition 4 of TARK is exactly the same as our Proposition 16. Note that the formalization of voting concepts in Definition 13 of TARK is very different from the formalization in our Section 7, both in form and in meaning: the propositional letters in the TARK logic stand for 'agent i prefers candidate a over candidate b', whereas our proposition letters stand for profiles.

[12] V. Conitzer, T. Walsh and L. Xia, 2011, "Dominating manipulations in voting with partial information", *Proc. of 25th AAAI*.

[13] H. van Ditmarsch, W. van der Hoek and B. Kooi, 2008, *Dynamic Epistemic Logic*, Berlin: Springer.

[14] H. van Ditmarsch, J. Lang and A. Saffidine, 2013, "Strategic voting and the logic of knowledge", *Proc. of 14th TARK – Chennai*, http://arxiv.org/abs/1310.6436.

[15] J. Duggan and T. Schwartz, 2000, "Strategic manipulability without resoluteness or shared beliefs: Gibbard-Satterthwaite generalized", *Social Choice and Welfare*, **17(1)**: 85–93.

[16] R. Fagin et al., 1995, *Reasoning about Knowledge*, Cambridge, MA: MIT Press.

[17] A. Gibbard, 1973, "Manipulation of voting schemes: A general result", *Econometrica*, **41(4)**: 587–601.

[18] J. Harsanyi, 1967–1968, "Games with incomplete information played by 'Bayesian' players, Parts I, II, and III", *Management Science*, **14(3,5,7)**: 159–182, 320–334, 486–502.

[19] N. Hazon et al., 2008, "Evaluation of election outcomes under uncertainty", *Proc. of 7th AAMAS*, pp. 959–966.

[20] W. Jamroga and W. van der Hoek, 2004, "Agents that know how to play", *Fundamenta Informaticae*, **63(2-3)**: 185–219.

[21] K. Konczak and J. Lang, 2005, "Voting procedures with incomplete preferences", *Proc. IJCAI-05 Multidisciplinary Workshop on Advances in Preference Handling*.

[22] A. Lomuscio, 1999, *Knowledge Sharing among Ideal Agents*, phdthesis, University of Birmingham.

[23] R. Meir, 2015, "Plurality voting under uncertainty", *Proc. of the 29th IJCAI*, pp. 2103–2109.

[24] R. Parikh, Ç. Taşdemir and A. Witzel, 2013, "The power of knowledge in games", *International Game Theory Review*, **15(04)**: 1340030.

[25] T. Perkov, 2016, "Natural deduction for modal logic of judgment aggregation", *Journal of Logic, Language and Information*, **25(3-4)**: 335–354.

[26] J. Plaza, 1989, "Logics of public communications", *Proc. of the 4th ISMIS*, pp. 201–216, Oak Ridge National Laboratory.

[27] A. Reijngoud and U. Endriss, 2012, "Voter response to iterated poll information", in W. van der Hoek et al. (eds.), *Proc. of 11th AAMAS*, pp. 635–644.

[28] M. A. Satterthwaite, 1975, "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.

[29] N. Troquard, W. van der Hoek and M. Wooldridge, 2011, "Reasoning about social choice functions", *Journal of Philosophical Logic*, **40(4)**: 473–498.

不确定性对策略性投票的影响

泽纳布·巴赫蒂亚尔　　　汉斯·范·狄马赫　　　阿卜杜拉·萨菲丁

摘　　要

　　如果选民不确定其他选民的偏好以及其他选民对他本人偏好的知晓程度，选票操纵的策略会变得更复杂。为了研究这种认知不确定性对策略性投票的影响，本文提供了一个形式分析框架并定义了选票操纵、优势偏好、均衡票况等概念。我们发现：即使真实偏好相同但关于偏好的高阶知识不同也会影响均衡票况。另外，本文还研究了不确定性的更新对选票操纵的影响，如某位选民对他本人的偏好进行公开宣告。我们证明了有些形式的选票操纵和均衡票况可以在这样的更新下得到保持，而另一些则不能被保持。最后，我们用一个逻辑语言形式化了不确定性选举的相关概念。本文的目的是给出一个分析和设计面向不确定性的投票机制的技术基础。

泽纳布·巴赫蒂亚尔	洛林计算机及其应用实验室，洛林大学及法国国家科学研究中心 bakhtiarizeinab@gmail.com
汉斯·范·狄马赫	洛林计算机及其应用实验室，洛林大学及法国国家科学研究中心 hans.van-ditmarsch@loria.fr
阿卜杜拉·萨菲丁	新南威尔士大学计算机科学与工程学院 abdallah.saffidine@gmail.com