Simultaneous manipulation under incomplete information

Yuliya A. Veselova and Daniel Karabekyan

Abstract

We study individual manipulability of social choice rules under incomplete information modeled with poll information functions (PIFs). Voters that have an incentive to manipulate under PIF π are called π-manipulators. Since voters do not know preferences of others exactly, there is an uncertainty about a situation they are in. In addition, we compare manipulability for different assumptions about voters' behavior which constitute a behavioral model. In the first model incentives to manipulation do not depend on other voters' incentives. In the second model each π-manipulator takes into account that all other π-manipulators strategize, and in the third model only a subset of them do so. In theoretical part, we study when non-manipulability is inherited from one model to another and reveal conditions for non-manipulability of scoring rules. With the help of computations we reveal how the type of information and behavioral model influence the relative manipulability of 12 social choice rules.

1 Introduction

When a group of individuals needs to make a collective decision, they can collect all individual preferences and process them with some aggregating procedure. However, voters can achieve a better voting result by misrepresenting their preferences, i.e. manipulate. It turned out that the class of voting procedures that allow for manipulations is very broad. Gibbard [13] and Satterthwaite [32] independently proved that in the presence of at least 3 candidates (or alternatives) any onto and non-dictatorial voting rule is vulnerable to manipulation. Different rules may be manipulable to a different extent. For this reason the most used approach to comparing the degree of manipulability of rules is measuring the share of situations (preference profiles) that admit manipulation by voters. This approach was first used by [24] and [17] and further applied to the analysis of a big variety of rules in different models [1, 18, 30, 4, 2, 3, 29].

In the majority of publications considering the probability of manipulation there are two important simplifying assumptions made: a) when deciding whether to manipulate or not voters do not take into account possible actions of other voters (so-called ”naive” behavior) ; b) voters possess all information about preferences of each other. The aim of this research is to change both assumptions and study manipulability of voting rules for different behavioral models and types of information. To deal with the first one we use a version of ”safe” strategy used by Slinko and White [35]. A voter having an incentive to manipulate individually with some strategy is called a Gibbard-Satterthwaite manipulator, or GS-manipulator. Apart from the basic model with naive voters (Model 1) we consider two non-naive models of voters' beliefs. In Model 2 each GS-manipulator considers what the result of manipulation will be if all other GS-manipulators act strategically. If it occurs that voting sincerely is better than manipulation for a voter provided that all other GS-manipulators do manipulate, then manipulation becomes risky and the voter looses an incentive to manipulate. In Model 3 voters believe that any subset of potential manipulators can actually act strategically, while others vote sincerely. This creates a higher level of uncertainty. And again if there is a risk of getting worse off by using a manipulation strategy instead of voting sincerely, a voter loses an incentive to manipulate.

Moreover, we add an assumption of incomplete information. Some models of manipulation under incomplete information have already been considered. We use the one presented by Reijngoud and Endriss [31]. It is assumed that all voters report their sincere preferences to an opinion poll held before voting. Then results of the opinion poll are made public. Since they are represented in an aggregated form (a result of a poll information function, PIF), voters do not know exactly each others' preferences. Thus, a voter has an incentive to manipulate under a given PIF if there is a strategy such that a voter has a chance of getting better off and has no chance of getting worse off with this strategy. We consider three types of PIF. The first one is the unique winner, the result of a voting rule after tie-breaking. The second one is the set of winners according to a voting rule. And the third is the ranking of alternatives (a weak order) produced by a voting rule.

Thus, combining the uncertainty from incomplete information with the uncertainty about other manipulators' actions we get a serious obstacle for a potentially manipulating voter. However, if a voter has an incentive to manipulate despite all these difficulties, this adds to manipulability of a voting rule. In the literature there are not many papers studying the probability of manipulation with non-naive voters and with incomplete information separately (see Section 2). This work is the first one (to the best of our knowledge) considering both and their mutual influence. We conduct computational experiments calculating exact manipulability indexes for different combination of information types and voters' beliefs about others. Moreover, we prove that for any number of alternatives there is a specific number of voters such that for any greater number of voters manipulability is impossible for any scoring rule when voters have information about a unique winner of an election if voters take into account other manipulators' actions.

2 Related literature

In this section we aim to consider more thoroughly the body of research devoted to interaction of voters, voting games, and informational aspects of voting for better understanding the place of our work in the literature.

If there is only one GS-manipulator among voters, then she is a pivotal voter and has an opportunity to influence a voting result on her own. If in a society there are several GS-manipulators, the voting result is difficult to predict due to the problem of their interaction. If a voter knows that other GS-manipulators may also decide to act strategically, can this affect her incentives to manipulation? This question was first considered in [35] where each GS-manipulator considers the possibility that other voters with the same preferences (and, consequently, also being GS-manipulators) may strategise. These authors define a strategy to be "safe" if regardless of what subset of other co-minded agents manipulates there is no possibility for a voter to become worse off and for at least one subset she becomes better off. This direction was followed by Hazon and Elkind [15] and Ianovski et al. [16] who studied computational complexity of finding a safe manipulative vote. The asymptotic probability of a safely manipulable profile for scoring rules was considered by Wilson and Reyhani Shokat Abad [40].

The next step for a voter is to think not only about her allies, but also about other people who have an incentive to manipulate. So, an extension of this model considers all GS-manipulators as players in a voting game. Then a strategy chosen by a manipulating voter can be called "safe" if it is at least as good as sincere voting for any possible actions of other GS-manipulators. For simplicity it is usually assumed that each manipulator chooses between truth-telling and one strategy chosen according to some optimality principle. This kind of model was considered in [10] and [14]. In these publications the existence of pure strategy Nash equilibria is studied for plurality and k-Approval voting rule with $k = 2,3,4$. Thus, this model of voters' behavior is the closest to Model 3 in our study. We also consider only GS-manipulators who choose between sincere voting and one manipulation strategy.

However, the set of players may not be restricted to the set of GS-manipulators. Voters which do not have an incentive to manipulate on their own may also be considered as players and pose a counter-threat to manipulators' actions. Pattanaik [27, 28] and Barberà [5] study how coalitions of voters could counter-manipulate in response to individual manipulations and influence incentives of GS-manipulator, and the game between manipulator and counter-manipulator was considered by Grandi et al. [14].

And what if the set of players is the whole set of voters and the set of their strategies is not restricted? This framework is the most general and was considered many times, for example, by Moulin [21] and Myerson and Weber [22] among the first. Both papers used Nash equilibrium as a solution concept, but faced the problem of a great multiplicity of equilibria. Since lots of these equilibria are weird, there appeared many papers suggesting different ways to eliminate them (see surveys by Meir [19] or Slinko [34]). For example, one way is to assume that voters prefer to abstain or to vote sincerely when they are not pivotal [9, 25]. Another one is to refine the set of Nash equilibria [8, 33, 9, 41, 26]. Moreover, it is possible to assume bounded rationality of voters, who may not think of other voters being strategic - and we come again to the aforementioned works of Slinko and White [35], Elkind et al. [10] and Grandi et al. [14].

A topic which follows directly from the previous one is modeling voter levels of rationality. In the structural level-k models of Nagel [23] and Stahl and Wilson [37] voters of level k of rationality believe that other voters are of level $k-1$. Thus, voters of level 0 do not strategize, voters of level 1 choose their best strategy in assumption that all other voters are of level 0, level 2 voters choose the best response believing that other voters are of level 1. A cognitive hierarchy (CH) model of Camerer et al. [6] has a difference that level-k voters believe that others can have any level from 0 to $k-1$. The CH-model was used in the work by Elkind et al. [11] which focuses on computational complexity of deciding whether a manipulation strategy weakly dominates a sincere vote for a level-2 voter. The CH-model is also applicable in our work. As in the work by Elkind et al. [11], we consider only the first three levels and assume that all voters not being GS-manipulators are of level 0.

A situation when voters do not know anything about actions of other voters, i.e. any voter can potentially submit any preference order, is equivalent to the zero-information case. Can a voter choose a strategy weakly dominating sincere voting in such a situation? Intuitively, no. Indeed, Moulin [21] mentioned that for voters having no information about other voters' preferences the best strategy is to vote sincerely. For Condorcet-consistent rules and Borda rule it was formally proved by Conitzer et al. [7], and for non-manipulability of scoring rules authors give the bound which was strengthened in the work by [31]. Ho. However, this is an extreme case and it seems more natural to assume that voters can predict actions of others to some extent (if they know their true preferences, like in the models mentioned above) or know something about preferences of a society (incomplete information).

Models of manipulation under incomplete information attract more attention in recent years. One of the first formal models for strategic voting with partial information was introduced by Conitzer et al. [7]. The main focus of the paper was complexity of manipulation. A similar model was introduced by Reijngoud and Endriss [31], but instead of partial orders, authors consider results of preelection opinion polls. This model was used for the analysis of manipulability of voting rules under various public information types by Veselova [39]. However, the main application sphere of opinion polls models is iterative voting, where they serve as a coordination device for voters [22, 31, 12, 20]. More complex models of information may include not only knowledge of other voters' preferences, but also knowledge about knowledge. For this purpose epistemic logic is used [38, 36].

Thus, the current work considers the model of individual manipulation by voters with bounded rationality under incomplete information. Each possible preference profile creates a game with GS-manipulators as players. As level 2 players in the work by Nagel [23] and Stahl and Wilson [37], in our behavioral Model 2 each manipulating agent thinks about others GS-manipulators as being level 1. And in Model 3 they admit that other GS-manipulators may be level 1 or 0 as in CH-models [6, 11]. In contrast to the mentioned works level 2 voters do not search for the best reply, but check whether their strategy for GS-manipulation under uncertainty still works when we add uncertainty about other voters' actions.

3 The Framework

3.1 The Model

Let $N = \{1,...,n\}$ be a set of *voters* which have preferences over a set of *alternatives* X, $|X| = m$. $P_i \subseteq X \times X$ is a *preference order* of agent i and it is assumed to be a linear order, i.e. irreflexive, weakly complete and transitive binary relation on X. A *preference profile* of all voters is denoted by $\mathbf{P} = (P_1,...,P_i,...,P_n)$ and a preference profile of all voters except i is $\mathbf{P}_{-i}$. The set of all preference profiles is $L(X)^N$ and includes $(m!)^n$ elements. A mapping $F : L(X)^N \to 2^X \setminus \emptyset$ is called a *voting rule*. If the result of the voting rule contains more than one alternative, then a *tie-breaking rule* (TBR) is used, $T : 2^X \setminus \emptyset \to X$. We use alphabetic tie-breaking: let some linear order on X to be predefined, for example, $aP_TbP_Tc...$, and when alternatives are tied, we choose the one which dominates all others by P_T, i.e $T(A) = \{a \in A | \forall x \in A, x \neq a\ (a,x) \in P_T\}$. The composition of functions F and T, i.e. $V(\mathbf{P})$ is denoted by $V(\mathbf{P})$. Although a result of a voting rule is a set of alternatives which are considered as the best ones, we could also define a ranking of alternatives based on this rule. Let us denote this ranking by a weak order P (irreflexive, transitive, and negatively transitive binary relation), an element of the set of all weak orders on X, $W(X)$.

By $v_j(a,\mathbf{P})$ we denote the number of voters having a on the j-th position in preferences (the most preferred alternative gets the 1st position). A vector of positions for an alternative a is $v(a,\mathbf{P}) = (v_1(a,\mathbf{P}),...,v_m(a,\mathbf{P}))$. By μ we denote *majority relation*: $a_k\mu a_l$ if $|\{i \in N : a_kP_ia_l\}| > |\{i \in N : a_lP_ia_k\}|$. A *matrix of a majority graph* is $MG(\mathbf{P})$, where $MG(\mathbf{P})_{kl} = 1$ if $a_k\mu a_l$, $MG(\mathbf{P})_{kl} = -1$ if $a_l\mu a_k$,$MG(\mathbf{P})_{kl} = 0$ otherwise.

Similar to Reijngoud and Endriss [31], Veselova [39], Endriss et al. [12], we use the *poll information function* $\pi(\mathbf{P})$ (PIF) that shows what kind of information is known by a voter about $\mathbf{P}$. In this paper, we consider 4 types of PIFs.

1. *1Winner*. Information only about the unique winner after the TBR, $\pi_{1Winner}(\mathbf{P}) = V(\mathbf{P})$
2. *Winner*. Information only about the winner(s) before tie-breaking, $\pi_{Winner}(\mathbf{P}) = F(\mathbf{P})$
3. *Rank*. Information about the ranking of alternatives, $\pi_{Rank}(\mathbf{P}) = R$.
4. *Profile*. Information about a full profile is known. It is the classic case of complete information. $\pi_{Profile}(\mathbf{P}) = \mathbf{P}$.

Let $W_i^{\pi(\mathbf{P})}$ be the *information set* of voter i, the set of all possible preference profiles of other voters that are consistent with information $\pi(\mathbf{P})$ of voter i.

$$W_i^{\pi(\mathbf{P})} = \{\mathbf{P}'_{-i} \in L(X)^{N\setminus\{i\}} : \pi(P_i,\mathbf{P}'_{-i}) = \pi(\mathbf{P})\}. \quad (1)$$

We say that π is *at least as informative as* π' if for all $\mathbf{P} \in L(X)^N$ and for all $i \in N$ we have $W_i^{\pi(\mathbf{P})} \subseteq W_i^{\pi'(\mathbf{P})}$. In our list of PIFs, they go from the least informative (1Winners-PIF) to the most informative (Profile-PIF).

Definition 1. *We say that a voter has an incentive to π-manipulate in a preference profile $\boldsymbol{P}$ under a rule F if there exists some preference $\tilde{P}_i$ such that i) either $V(\tilde{P}_i,\boldsymbol{P}'_{-i}) = V(\boldsymbol{P})$ or $V(\tilde{P}_i,\boldsymbol{P}'_{-i})P_iV(\boldsymbol{P})$ for all $\boldsymbol{P}'_{-i}$ in $W_i^{\pi(\boldsymbol{P})}$ and ii) $V(\tilde{P}_i,\boldsymbol{P}'_{-i})P_iV(\boldsymbol{P})$ for at least one $\boldsymbol{P}'$ in $W_i^{\pi(\boldsymbol{P})}$.*

In other words, a voter will manipulate if for every possible profile of her information set she gets at least the same result and for at least one profile she gets a more preferable alternative, on the condition that all others vote sincerely.[1]

If at least one voter has an incentive to π-manipulate in $\mathbf{P}$ under F, then preference profile $\mathbf{P}$ is π-manipulable under F. A voter having an incentive to π-manipulate is called a *π-manipulator*. The set of all π-manipulators in profile $\mathbf{P}$ is denoted by $\Pi(\mathbf{P})$.

In the Definition 1 proposed by Reijngoud and Endriss [31] a voter does not think about possible actions of others. We would like to consider other assumptions on voters' behavior and formalize this in the term "behavioral model". We consider three behavioral models: Model 1 suggests that voters behavior is naive, they do not think about actions of others (definition of manipulation is the same as the basic one); in Model 2 voters check whether their manipulation strategy still works when all other π-manipulators manipulate as well; in Model 3 voters think that there could be some other voters manipulating.

3.2 Voting rules

- *Scoring rules*. A scoring rule is defined by a scoring vector $s = (s_1, ..., s_m)$, where s_j denotes the score assigned to an alternative at the j-th position in individual preferences. The total score of an alternative a is denoted by $S(a,\mathbf{P})$ and calculated as a dot product $s \cdot v(a_j,\mathbf{P}) = \sum_{i=1}^{m} s_i v_i(a_j,\mathbf{P})$. Then a social ordering P is defined as follows: for all $a,b \in X$ i) $aPb \Leftrightarrow S(a,\mathbf{P}) > S(b,\mathbf{P})$. Alternatives with maximum score win, i.e. $x \in F(\mathbf{P}) \Leftrightarrow x \in argmax_{a\in X} S(a,\mathbf{P})$.
  - *Plurality*: $s_{Pl} = (1,0,...,0)$.
  - *Veto* (Antiplurality): $s_V = (1,...,1,0)$.
  - *Borda*: $s_B = (m-1, m-2, ..., 1, 0)$.
- *Run-off procedure*. It has two stages:
  1) For each alternative $x \in X$ the plurality score is calculated, $S^1(x,\mathbf{P}) = s_{Pl} \cdot v(x,\mathbf{P})$. If $\exists a \in X$ such that $S^1(a,\mathbf{P}) > n/2$, then for all $b \in X \setminus \{a\}$, aPb, and procedure terminates. Otherwise, procedure moves on to the stage two.
  2) Two alternatives with maximal number of scores are chosen. If there are ties, they are broken according to the tie-breaking rule T.
  $$a = T(argmax_{x\in X} S^1(x,\mathbf{P})), \tag{2}$$
  $$b = T(argmax_{x\in X\setminus\{a\}} S^1(x,\mathbf{P})). \tag{3}$$
  The chosen alternatives dominate all other alternatives in ranking P: for all $x \in X \setminus \{a,b\}$ aPx, bPx. Finally, if $a\mu b$, then aPb and $F(\mathbf{P}) = \{a\}$ if $b\mu a$, then bPa and $F(\mathbf{P}) = \{b\}$, and if $|\{i \in N : aP_ib\}| = |\{i \in N : bP_ia\}|$, then $F(\mathbf{P}) = \{a,b\}$.
- *Single Transferable vote (STV)*. This is a multi-stage procedure, which we define in an iterative form.
  0) $t := 1$, $X^t := X$, $\mathbf{P}^t := \mathbf{P}$.
  1) For all $a \in X^t$ $S^t(a,\mathbf{P}) := s_{Pl} \cdot v(a,\mathbf{P}^t)$.
  2) If there exists $a \in X^t$ such that $S^t(a,\mathbf{P}) > n/2$, then $F(\mathbf{P}) = \{a\}$ and the procedure terminates $\forall b \in X^t \setminus \{a\}$ aPb. Else $A := argmin_{a\in X^t}(S^t(a,\mathbf{P}))$.
  3) If $A = X^t$, then $F(\mathbf{P}) = X^t$ and the procedure terminates. Otherwise, alternatives of A are eliminated, $t := t+1$, $X^t := X^{t-1} \setminus A$, $\mathbf{P}^t := \mathbf{P}/X^t$; for all $x \in X^t$ and $a \in A$ it holds xPa; go to step 1.

[1]One can imagine an alternative definition of manipulation with only successful outcomes: a voter manipulates only if her manipulation leads to a success for all preference profiles of her information set. However, in this case manipulation becomes a very rare event and even impossible for some settings.

- *Copeland*. A majority graph is computed. Then scores of alternatives are computed as follows

$$S(a_k, \mathbf{P}) = \sum_{l=1}^{m} MG(\mathbf{P})_{kl}. \tag{4}$$

  Ranking P is defined on the base of a scoring function: for all $a, b \in X$ $aPb \Leftrightarrow S(a, \mathbf{P}) > S(b, \mathbf{P})$. Alternatives with maximum score win, i.e. $x \in F(\mathbf{P}) \Leftrightarrow x \in argmax_{a \in X} S(a, \mathbf{P})$.
- *Maximin*. For each alternative $x \in X$ the number of scores is computed as follows $S(x, \mathbf{P}) = min_{a \in X} |\{i \in N : xP_i a\}|$. Then alternatives are ranked according to the number of scores: for all $a, b \in X$ i) $aPb \Leftrightarrow S(a, \mathbf{P}) > S(b, \mathbf{P})$. Alternatives with maximum score win, i.e. $x \in F(\mathbf{P}) \Leftrightarrow x \in argmax_{a \in X} S(a, \mathbf{P})$.
- *Baldwin's rule*. Multistage procedure.
  0) $t := 1$, $X^t := X$, $\mathbf{P}^t := \mathbf{P}$.
  1) For all $a \in X^t$ count Borda score $S^t(a, \mathbf{P}^t) := s_B \cdot v(a, \mathbf{P}^t)$.
  2) Find alternatives with minimum score $A := argmin_{a \in X^t}(S^t(a, \mathbf{P}))$.
  3) If $A = X^t$, then $F(\mathbf{P}) = X^t$ and the procedure terminates. Otherwise, alternatives of A are eliminated, $t := t + 1$, $X^t := X^{t-1} \setminus A$, $\mathbf{P}^t := \mathbf{P}/X^t$; for all $x \in X^t$ and $a \in A$ it holds xPa; go to step 1.
- *Nanson's rule*. For each alternative
  0) $t := 1$, $X^t := X$, $\mathbf{P}^t := \mathbf{P}$.
  1) For all $a \in X^t$ $S^t(a, \mathbf{P}^t) := s_B \cdot v(a, \mathbf{P}^t)$.
  2) Compute the average score

$$\bar{r}^t = \sum_{a \in X^t} S^t(a, \mathbf{P}^t)/|X^t|. \tag{5}$$

  3) Find alternatives that have the score lower than $\bar{r}^t$: $A := \{a \in X^t | S^t(a, \mathbf{P}) < \bar{r}^t\}$.
  4) If A is empty, then $F(\mathbf{P}) = X^t$ and the procedure terminates. Otherwise, alternatives of A are eliminated, $t := t + 1$, $X^t := X^{t-1} \setminus A$, $\mathbf{P}^t := \mathbf{P}/X^t$; for all $x \in X^t$ and $a \in A$ it holds xPa; go to step 1.
- *Black' procedure*. Procedure chooses a Condorcet winner $CW(\mathbf{P}) = [a | \neg\exists x \in X, x\mu a]$ if it exists, and then for all $x \in X$ $CW(\mathbf{P})Px$. Otherwise, the Borda rule is applied.
- *Kemeny's rule*. Let the distance between linear orders be a function $d(P_i, P_j) = |(P_i \setminus P_j) \cup (P_j \setminus P_i)|$. then P is an ordering such that $P = argmin_{R \in L(X)} \sum_{i \in N} d(R, P_i)$. The top alternative of P is the winner.
- *Threshold rule*. Alternatives are ordered on the basis of their position vectors: aPb if $v_m(a, \mathbf{P}) < v_m(b, \mathbf{P})$, or if there exist $k \le m$ such that $v_i(a, \mathbf{P}) = v_i(a, \mathbf{P})$, $i = k-1, ..., m$, and $v_k(a, \mathbf{P}) < v_k(b, \mathbf{P})$. In words, we compare the number of worst positions of alternatives. If they are equal, then we compare the number of second-worst positions, and so on. Undominated alternatives are winners: $x \in F(\mathbf{P}) \Leftrightarrow \neg\exists a \in X$, such that aPx.

4 Model 1: naive manipulation

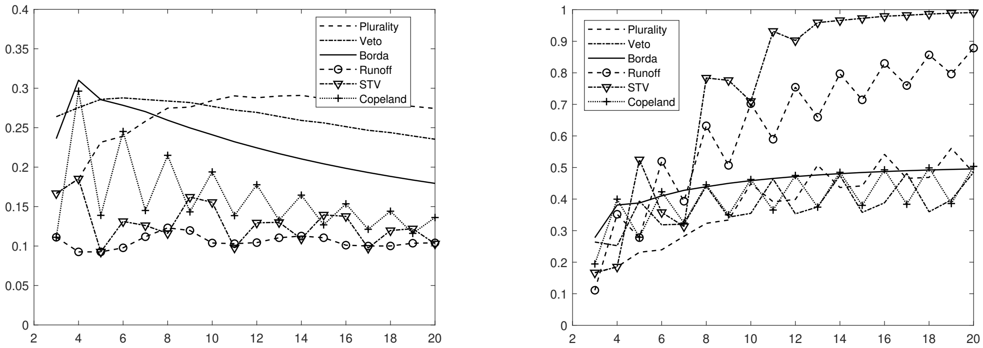
In this section we discuss the basic model of individual manipulation under incomplete information, which assumes that a voter does not think about incentives of other voters. Let us define it formally.

Definition 2. *A voter has an incentive to π-manipulate in Model 1 ($M1$) in $\mathbf{P}$ under F if and only if she has an incentive to π-manipulate in $\mathbf{P}$ under F.*

To compare the degree of manipulability of voting rules one needs some measure. Although in the case of complete information there exist different manipulability measures [39], we will use the simplest one, which is the proportion of preference profiles where manipulation is possible.

$I^M(m, n, \pi, F)$ – the share of preference profiles where at least one voter has an incentive to π-manipulate in model M under a rule F.

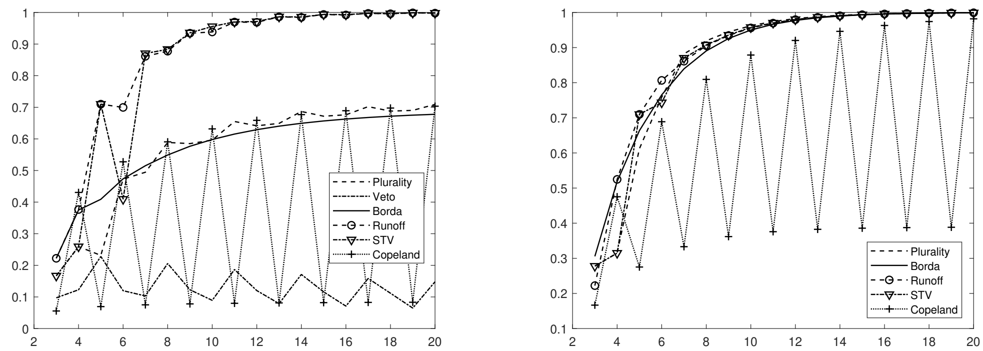
We computed the values of $I^{M1}(m,n,Profile,F)$ in MATLAB for $m=3$, and n from 3 to 20 changing the rule and the type of PIF. First we provide computational results for for the simplest model, individual naive manipulation with complete information, to compare them with further observations (see Fig.1). The graphs for the first 6 rules are shown in Figure 1 and Figure 2, for other rules - in Figure 3 and Figure 4 in Appendix A.



a) *Profile*-PIF　　　b) *Rank*-PIF

Figure 1: The values of $I^{M1}(3,n,\pi,F)$

As can be seen, all the computed values of I^{M1} are not greater than 0.4. For most rules the trend is slowly decreasing and among the least manipulable rules are runoff procedure, STV and Baldwin rule. Now change information type to *Rank*-PIF.



a) *Winner*-PIF　　　b) *1Winner*-PIF

Figure 2: The values of $I^{M1}(3,n,\pi,F)$

Lets us consider I^{M1} for *Rank*-PIF, Fig.1b (and Fig.3b in Appendix A). The less information is available to voters, the larger are voters' information sets. Profiles of the information set of voter i, $W_i^{\pi(\mathbf{P})}$, all give the same ranking when voter i does not manipulate. However, for all rules under consideration the result after manipulation is not always the same for all profiles of voter's information set $W_i^{\pi(\mathbf{P})}$. In other words, information about ranks of alternatives does not allow a voter to compute the result of manipulation for any strategy she chooses (using the term from [31], these rules are not strongly computable from Rank-images).

According to the definition, using a manipulation strategy must lead to a better result in some profiles of $W_i^{\pi(\mathbf{P})}$ and must not lead to a worse in others. The voter cannot distinguish between different profiles of her information set. So, if voter i has an incentive to manipulate in $\mathbf{P}$, then all profiles of $W_i^{\pi(\mathbf{P})}$ are manipulable, even those where voter i cannot really change anything. A reader can find a more detailed explanation of this effect in [39].

For most cases the values of I^{M1} for *Rank*-PIF are greater that for *Profile*-PIF. And the general trend for most rules changes to an increasing with near-1 values for STV, Baldwin, and Nanson rules and $m > 12$.

For *Winner*-PIF (Fig.2a and Fig.4a) even more rules show graphs going to 1 with growing n. These are the same as for *Rank*-PIF plus Black procedure, maximin, and runoff. Except for Copeland and veto rule, all other rules demonstrate values of $I^{(1)}$ for *Winner*-PIF greater than for *Rank*-PIF for almost all n. The graph for Copeland rule has a higher amplitude and manipulability of veto rule decreased considerably for all n.

For the least informative PIF, *1Winner*-PIF (Fig.2b and Fig.4b), all rules except for Copeland and veto, merge near 1 for $n > 6$. Peaks of the graph for Copeland rule also approach 1, and veto rule disappears from the figure due to the zero-manipulability for *1Winner*-PIF [31].

5 Model 2: manipulation with respect to all other π-manipulators

In this section we investigate what will change if a manipulating voter takes into account manipulations of others. In general, each π-manipulator has a set of strategies with that she has an incentive to π-manipulate. For simplicity, for each π-manipulator j in $\Pi(\mathbf{P}) \setminus \{i\}$ we fix one strategy $\hat{P}_j$ that she may use or not, which we will refer to as *π-manipulation strategy*.[2] In other words, voter i thinks that another voter j can π-manipulate only with a strategy $\hat{P}_j$. A π-manipulation strategy $\hat{P}_j$ is chosen by the principle of the best winning alternative and in case of equality we choose it alphabetically.

We denote by $\hat{\mathbf{P}}$ a preference profile obtained from $\mathbf{P}$ with the difference that all voters from $\Pi(\mathbf{P})$ use their π-manipulation strategy.

Definition 3. *We say that a voter i has an incentive to π-manipulate in Model 2 (M2) in a preference profile $\mathbf{P}$ under a rule F if there is a strategy $\tilde{P}_i$ such that:*
i) voter i has an incentive to π-manipulate with $\tilde{P}_i$ in $\mathbf{P}$ under F;
ii) it is either $V(\tilde{P}_i, \hat{\mathbf{P}}'_{-i}) = V(P_i, \hat{\mathbf{P}}'_{-i})$ or $V(\tilde{P}_i, \hat{\mathbf{P}}'_{-i}) P_i V(P_i, \hat{\mathbf{P}}'_{-i})$ for all $\mathbf{P}'_{-i}$ in $W_i^{\pi(\mathbf{P})}$.

In words, condition ii) requires that a strategy $\tilde{P}_i$ is still not worse than truth-telling provided that all other π-manipulators decide to manipulate. Thus, if some strategy $\tilde{P}_i$ for voter i fails to dominate truth-telling, we need to check another strategy, and so on. If none of them is dominant, then voter i does not have an incentive to π-manipulate in Model 2. So, although we fix manipulation strategies for other π-manipulators, we still need to consider all manipulation strategies for voter i. However, for the 3-alternatives case this occurs to be excessive, since different manipulation strategies do not differ in terms of the result.

The first series of experiments refers to the complete information case. However, instead of values $I^{M2}(3, n, Profile, F)$ we show results for $I^{M1}(3, n, Profile, F) - I^{M2}(3, n, Profile, F)$, which is more illustrative. Particularly, it shows in which proportion of naively manipulable preference profiles

[2] Otherwise, we should have considered all combinations of strategies for π-manipulators. This is not only computationally hard, but also not very interesting. For the 3-alternatives case if there are several manipulation strategies, then most likely they are equivalent (always give the same result).

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.06	.04	.03	.04	.03	.02	.03	.02	.02	.02	.02	.01	.01	.01	.01	.01	.01	.01
Veto	.06	.07	.08	.09	.09	.08	.08	.08	.07	.07	.07	.06	.06	.06	.05	.05	.05	.04
Borda	.00	.04	.03	.03	.02	.02	.01	.01	.01	.01	.01	.01	.01	.00	.00	.00	.00	.00

Table 1: $I^{M1}(3,n,Profile,F) - I^{M2}(3,n,Profile,F)$

(manipulable in M1) a threat of having something bad as the result in case of simultaneous manipulation destroys all incentives to manipulate. In this and the proceeding section computational experiment results are shown for scoring rules only. The full tables can be seen in Appendix B. As can be seen from Table 1 and Table 5 (Appendix B), this situation occurs most often under veto and threshold rule. There is almost no difference between Model 1 and Model 2 for runoff procedure, maximin, Baldwin and Kemeny rules under Profile-PIF. Except for veto and threshold rules, all other values do not exceed 0.07 and most of them are very close to 0. This suggests that having a threat to loose when all other manipulators do manipulate cannot be a serious obstacle to strategic voting, simply because it is quite rare.

Then we change information type to less and less informative and two opposite effects start to work together. On the one hand, the probability to meet a π-manipulable profile grows (see Fig.1-4) in general for many rules. On the other hand, the risk to result with something worse than initially due to simultaneous manipulation combined with uncertainty about preferences of others makes many profiles non-manipulable.

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality (*Rank*-PIF)	.00	.07	.12	.12	.18	.23	.25	.36	.31	.33	.43	.37	.39	.47	.40	.42	.50	.42
Veto (*Rank*-PIF)	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda (*Rank*-PIF)	.28	.19	.08	.08	.07	.06	.07	.08	.09	.10	.10	.11	.11	.12	.12	.12	.13	.13
Plurality (*Winner*-PIF)	.00	.10	.12	.14	.26	.28	.25	.33	.33	.30	.37	.36	.34	.39	.37	.36	.40	.38
Veto (*Winner*-PIF)	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda (*Winner*-PIF)	.22	.31	.15	.15	.16	.17	.19	.20	.22	.23	.24	.25	.26	.27	.27	.28	.28	.29
Plurality (*1Winner*-PIF)	.00	.15	.00	.17	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Veto (*1Winner*-PIF)	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda (*1Winner*-PIF)	.31	.27	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 2: $I^{M2}(3,n,PIF,F)$ for 3 rules and 3 PIFs

Let us list the most important observations from Table 2 and put them into groups.

Rank-PIF:

- Veto rule is non-manipulable for $n \geq 4$ and Nanson's is non-manipulable for $n \geq 11$;
- Borda and Maximin rules are the least manipulable;
- Plurality, Kemeny, and threshold rules are the most manipulable.

Winner-PIF:

- STV rule is non-manipulable for $n \geq 13$;
- For Copeland, maximin, Baldwin, Black and Kemeny rules there is an alternation of zero and non-zero values of manipulability index;
- Plurality, Borda, and threshold rules are the most manipulable.

1Winner-PIF:

- For all rules except for Copeland rule there is a value n' such that all values of $I^{M2}(m,n,F,\pi)$ are zeros for all $n \geq n'$.

We also need to mention that the computed values for Kemeny and maximin rules coincide for all PIFs except *Rank*-PIF. The coincidence is explained by the same results these rules give for the 3-alternative case. And the difference for *Rank*-PIF is caused by the different ways the *Rank*-PIF is constructed for these rules.

From the tables it can be seen that if we get zero manipulability in some case and increase the level of uncertainty, then zero manipulability is preserved. For example, we fix a PIF and go from Model 1 to Model 2.

Proposition 1. *For any PIF π, for any rule F, the number of voters n, and the number of alternatives m if $I^{M1}(m,n,F,\pi') = 0$, then $I^{M2}(m,n,F,\pi'') = 0$.*

The main observation for Model 2 and *1Winner*-PIF is that for all rules except Copeland manipulability indexes become 0 when the number of voters exceeds a certain value. It turns out that the same holds for any given number of alternatives for any scoring rule. First we prove Lemma 1 and then use it in Theorem 1 stating this.

For a scoring vector s, a *jump* is a non-zero difference between two adjacent scoring values. If s has r jumps, then this means that there are distinct $k_1,\dots,k_r \in \{1,\dots,m-1\}$ such that $s_{k_1} - s_{k_1+1} > 0,\dots,s_{k_r} - s_{k_r+1} > 0$, while all other differences are zero. Let $\Delta_j = s_{k_j} - s_{k_j+1}$ for $j = 1,\dots,r$ denote the j-th jump.

Lemma 1. *For any scoring rule F, any number of alternatives m, any voter i, any jump Δ_j, and any two distinct alternatives a_h and a_l there is a number of voters n^*, such that for all $n > n^*$ there exists $\boldsymbol{P} \in L(X)^N$ such that*
1) $P_i = (a_1, a_2, ..., a_m)$;
2) for all $a_g \in X \setminus \{a_h, a_l\}$ $S(a_h, \boldsymbol{P}) > S(a_g, \boldsymbol{P})$;
3) $S(a_h, \boldsymbol{P}) - S(a_l, \boldsymbol{P}) = \Delta_j$ ($S(a_h, \boldsymbol{P}) - S(a_l, \boldsymbol{P}) = 0$).

Theorem 1. *For any scoring rule F and any number of alternatives m there is a finite number of voters n^*, such that for all $n > n^*$ it holds $I^{M2}(m,n,F,1Winner) = 0$.*

The proofs of Lemma 1 and Theorem 1 can be found in Appendix C.

6 Model 3: manipulating subsets of GS-manipulators

In this model, a voter assumes that each of the other π-manipulators may manipulate or not. Thus, it is necessary to consider all possible subsets K of $\Pi(\mathbf{P}) \setminus \{i\}$. Then voter i compares the result with using strategy $\tilde{P}_i$ and voting sincerely provided that voters from K manipulate and others do not.

It is quite natural to assume that not all potential manipulators will actually manipulate. The problem is that for a voter it is computationally hard to think about actions of all possible subsets of π-manipulators. We show, however, that Model 3 does not differ very much from the previous one. Results are certainly strengthened, but not essentially different. Let us give a formal definition of manipulation in Model 3.

Definition 4. *We say that a voter i has an incentive to π-manipulate in Model 3 ($M3$) in a preference profile $\boldsymbol{P}$ under a rule F if there is a strategy $\tilde{P}_i$ such that:*
i) voter i has an incentive to π-manipulate with $\tilde{P}_i$ in $\boldsymbol{P}$ under F;
ii) it is either $V(\tilde{P}_i, (\hat{\boldsymbol{P}}'_K, \boldsymbol{P}'_{-K-i})) = V(P_i, (\hat{\boldsymbol{P}}'_K, \boldsymbol{P}'_{-K-i}))$ or
$V(\tilde{P}_i, (\hat{\boldsymbol{P}}'_K, \boldsymbol{P}'_{-K-i})) P_i V(P_i, (\hat{\boldsymbol{P}}'_K, \boldsymbol{P}'_{-K-i}))$ for all $\boldsymbol{P}'_{-i}$ in W_i^π and for all $K \subseteq \Pi(\boldsymbol{P}') \setminus \{i\}$.

Analogously, condition ii) requires that a strategy $\tilde{P}_i$ is still not worse truth-telling if some of π-manipulators also decide to manipulate.

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.00	.02	.03	.02	.03	.04	.03	.04	.04	.03	.04	.04	.03	.03	.03	.03	.03
Veto	.00	.01	.02	.03	.03	.05	.06	.07	.08	.09	.09	.10	.11	.11	.12	.12	.13	.13
Borda	.00	.00	.01	.03	.03	.05	.05	.06	.06	.07	.07	.07	.07	.07	.08	.08	.08	.08

Table 3: $I^{M2}(3,n,Profile,F) - I^{M3}(3,n,Profile,F)$

Again, let us first consider the difference of manipulability indexes in Model 2 and Model 3 for the complete information case (Table 3 and Table 9 in Appendix B). On average, this difference has higher values than the difference between Model 1 and Model 2. This shows that the number of profiles where a subset of manipulators can spoil the result by their strategic actions is larger than the number of profiles where this can be done by all manipulators.

High values are more likely for bigger numbers of voters, since the number of GS-manipulators is greater in this case. The highest values of difference correspond to veto, Borda, Nanson's and threshold rules. And the least difference show maximin, Kemeny (which are the same again), and Baldwin rules. Since Model 3 assumes even higher level of uncertainty for voters than Model 2, zero-manipulability obtained for Model 2 is inherited by Model 3. Let us formulate this inheritance in the following two propositions.

Proposition 2. *For any PIF π, for any rule F, the number of voters n, and the number of alternatives m if $I^{M1}(m,n,\pi,F) = 0$, then $I^{M3}(m,n,\pi,F) = 0$.*

Proposition 3. *For any PIF π, for any rule F, the number of voters n, and the number of alternatives m if $I^{M2}(m,n,\pi,F) = 0$, then $I^{M3}(m,n,\pi,F) = 0$.*

Finally, if we fix the behavioral model and go from a more informative PIF π' to a less informative PIF π'', then 0-manipulability obtained for π' is preserved for π''.

Proposition 4. *Suppose, π' is at least as informative as π''. Then for any $M \in \{M1, M2, M3\}$, for any rule F, number of voters n, and number of alternatives m if $I^M(m,n,\pi',F) = 0$, then $I^M(m,n,\pi'',F) = 0$.*

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality (*Rank*-PIF)	.00	.07	.12	.12	.18	.23	.25	.36	.31	.33	.43	.37	.39	.47	.40	.42	.50	.42
Veto (*Rank*-PIF)	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda (*Rank*-PIF)	.28	.19	.08	.08	.07	.06	.07	.08	.09	.10	.10	.11	.11	.12	.12	.12	.13	.13
Plurality (*Winner*-PIF)	.00	.10	.12	.14	.26	.28	.25	.33	.33	.30	.37	.36	.34	.39	.37	.36	.40	.38
Veto (*Winner*-PIF)	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda (*Winner*-PIF)	.22	.31	.15	.15	.16	.17	.19	.20	.22	.23	.24	.25	.26	.27	.27	.28	.28	.29
Plurality (*1Winner*-PIF)	.00	.15	.00	.17	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Veto (*1Winner*-PIF)	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda (*1Winner*-PIF)	.31	.27	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 4: $I^{M3}(3,n,PIF,F)$ for 3 rules and 3 PIFs

In Table 4 (and Tables 10-12 in Appendix B) zero values obtained for Model 2 are highlighted with a darker grey, and those for Model 3 with a lighter grey. First, let us look at the top part of Table 4, which shows the manipulability index for *Rank*-PIF and behavioral Model 3. It can be seen that only veto and Nanson's rule demonstrate clear zero-tending dynamics of $I^{M3}(m,n,F,\pi)$, and values for maximin rule and Black procedure are quite low. Zero-manipulability occurs in some cases for runoff, STV, and maximin. However, most of zero values are inherited from Model 2. Plurality, Kemeny, and threshold rules are the most manipulable.

Now let us move on to the middle part of Table 4 (and Table 11 in Appendix B), corresponding to *Winner*-PIF. For plurality, Borda, and threshold rules if voters know information about winners before tie-breaking, they still can manipulate, even when they think that others may also manipulate. We view this as an important observation, since it puts these rules in contrast with the remaining ones, which give non-manipulability or zero manipulability alternating with non-zero manipulability from certain number of voters.

A slight difference in the quality of public information, i.e. voters do not know which alternatives were tied (in case there is a tie) and know only a final winner – this difference makes almost all rules impossible to manipulate from a certain value of n. Periodicity is preserved for Copeland rule, and the amplitude seems to be growing. It means that Copeland rules provides an opportunity to manipulate even under uncertainty of preferences and intentions of others when the number of voters is even. However, other rules give us a hope that manipulation actually does not make much sense when you know little.

7 Conclusion

We have reviewed many works devoted to the problem of manipulability of voting rules. They considered the problem from probabilistic, computational, informational points of view. These aspects were considered separately and results are often unfavorable: rules turn out to be manipulable (Gibbard-Satterthwaite theorem type), and, moreover, highly manipulable (computed probabilities) or efficiently manipulable. The current study is the first one combining informational and behavioral aspect of manipulation and applying a probabilistic approach to it.

One may argue that computing all possible situations consistent with public information and predicting other manipulators' actions is too hard for a voter. Of course, it is, and this adds to the immunity of rules to manipulation in reality. However, the goal of this research was not to show that manipulation is hard, but to demonstrate that it does not make much sense.

Another possible objection is that there are still many cases when manipulation is probable. The answer here is following: we considered only the probability of that a voter will have an incentive to manipulate, but did not compute the probability of a successful manipulation. As shown by Veselova [39], this probability turns out to be rather small.

The deeper analysis of the manipulation models from this study may also include computation and comparison of the probability of success and failure for each voter. This could help to understand whether a voter is assumed to be too risk-averse when she does not manipulate in a view of one unfavorable situation. However, considering only strategies that do not spoil the result whatever other voters do seems quite natural for modeling voters behavior. It guarantees that a voter will not loose. The fact revealed by this study that under very natural conditions it is impossible to find such strategy in a vast number of cases is definitely good news.

Acknowledgements The article was prepared within the framework of the HSE University Basic Research Program. The authors would like to thank Hans Peters and Ton Storcken for their comments and suggestions. We also thank anonymous referees for the valuable comments.

References

[1] Fuad Aleskerov and Eldeniz Kurbanov. Degree of manipulability of social choice procedures. In *Current trends in economics*, pages 13–27. Springer, 1999.

[2] Fuad Aleskerov, Daniel Karabekyan, M Remzi Sanver, and Vyacheslav Yakuba. On the degree of manipulability of multi-valued social choice rules. *Homo Oeconomicus*, 28, 2011.

[3] Fuad Aleskerov, Daniel Karabekyan, M Remzi Sanver, and Vyacheslav Yakuba. On the manipulability of voting rules: the case of 4 and 5 alternatives. *Mathematical Social Sciences*, 64 (1):67–73, 2012.

[4] Fuad T. Aleskerov, Daniel C. Karabekyan, Remzi Sanver, and Vyacheslav Yakuba. Evaluation of the degree of manipulability of known aggregation procedures under multiple choice. *Journal of the New Economic Association*, 1(1):37–61, 2009.

[5] Salvador Barberà. Stable voting schemes. *Journal of Economic Theory*, 23(2):267–274, 1980.

[6] Colin F Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A cognitive hierarchy model of games. *The Quarterly Journal of Economics*, 119(3):861–898, 2004.

[7] Vincent Conitzer, Toby Walsh, and Lirong Xia. Dominating manipulations in voting with partial information. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence Dominating*, volume 11, pages 638–643, 2011.

[8] Francesco De Sinopoli. Sophisticated voting and equilibrium refinements under plurality rule. *Social Choice and Welfare*, 17(4):655–672, 2000.

[9] Yvo Desmedt and Edith Elkind. Equilibria of plurality voting with abstentions. In *Proceedings of the 11th ACM conference on Electronic commerce*, pages 347–356, 2010.

[10] Edith Elkind, Umberto Grandi, Francesca Rossi, and Arkadii Slinko. Gibbard-satterthwaite games. In *IJCAI*, 2015.

[11] Edith Elkind, Umberto Grandi, Francesca Rossi, and Arkadii Slinko. Cognitive hierarchy and voting manipulation in k-approval voting. *Mathematical Social Sciences*, 108:193–205, 2020.

[12] Ulle Endriss, Svetlana Obraztsova, Maria Polukarov, and Jeffrey S Rosenschein. Strategic voting with incomplete information. AAAI Press/International Joint Conferences on Artificial Intelligence, 2016.

[13] Allan Gibbard. Manipulation of voting schemes: a general result. *Econometrica*, 41(4):587–601, 1973.

[14] Umberto Grandi, Daniel Hughes, Francesca Rossi, and Arkadii Slinko. Gibbard–satterthwaite games for k-approval voting rules. *Mathematical Social Sciences*, 99:24–35, 2019.

[15] Noam Hazon and Edith Elkind. Complexity of safe strategic voting. In *International Symposium on Algorithmic Game Theory*, pages 210–221. Springer, 2010.

[16] Egor Ianovski, Lan Yu, Edith Elkind, and Mark C Wilson. The complexity of safe manipulation under scoring rules. In *IJCAI*, volume 11, pages 246–251, 2011.

[17] Jerry S Kelly. 4. minimal manipulability and local strategy-proofness. *Social Choice and Welfare*, 5(1):81–85, 1988.

[18] Dominique Lepelley and Fabrice Valognes. Voting rules, manipulability and social homogeneity. *Public Choice*, 116(1-2):165–184, 2003.

[19] Reshef Meir. Strategic voting. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 13(1):1–167, 2018.

[20] Reshef Meir, Maria Polukarov, Jeffrey S Rosenschein, and Nicholas R Jennings. Iterative voting and acyclic games. *Artificial Intelligence*, 252:100–122, 2017.

[21] Hervé Moulin. Prudence versus sophistication in voting strategy. *Journal of Economic Theory*, 24(3):398–412, 1981.

[22] Roger B Myerson and Robert J Weber. A theory of voting equilibria. *American Political Science Review*, pages 102–114, 1993.

[23] Rosemarie Nagel. Unraveling in guessing games: An experimental study. *The American Economic Review*, 85(5):1313–1326, 1995.

[24] Shmuel Nitzan. The vulnerability of point-voting schemes to preference variation and strategic manipulation. *Public Choice*, 47(2):349–370, 1985.

[25] Svetlana Obraztsova, Evangelos Markakis, and David RM Thompson. Plurality voting with truth-biased agents. In *International Symposium on Algorithmic Game Theory*, volume 8146 LNCS of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, pages 26–37. Springer, 2013.

[26] Svetlana Obraztsova, Zinovi Rabinovich, Edith Elkind, Maria Polukarov, and Nicholas R Jennings. Trembling hand equilibria of plurality voting. pages 440–446. AAAI Press/International Joint Conferences on Artificial Intelligence, 2016.

[27] Prasanta K Pattanaik. Threats, counter-threats, and strategic voting. *Econometrica: Journal of the Econometric Society*, pages 91–103, 1976.

[28] Prasanta K Pattanaik. Counter-threats and strategic manipulation under voting schemes. *The Review of Economic Studies*, 43(1):11–18, 1976.

[29] Hans Peters, Souvik Roy, and Ton Storcken. On the manipulability of approval voting and related scoring rules. *Social Choice and Welfare*, 39(2-3):399–429, 2012.

[30] Geoffrey Pritchard and Mark C Wilson. Exact results on manipulability of positional voting rules. *Social Choice and Welfare*, 29(3):487, 2007.

[31] Annemieke Reijngoud and Ulle Endriss. Voter response to iterated poll information. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 635–644, 2012.

[32] Mark Allen Satterthwaite. Strategy-proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217, 1975.

[33] M. R. Sertel and M. R. Sanver. Strong equilibrium outcomes of voting games are the generalized condorcet winners. *Social Choice and Welfare*, 22(2):331–347, 2004.

[34] Arkadii Slinko. Beyond gibbard and satterthwaite: Voting manipulation games. In *The Future of Economic Design*, pages 131–138. Springer, 2019.

[35] Arkadii Slinko and Shaun White. Is it ever safe to vote strategically? *Social Choice and Welfare*, 43(2):403–427, 2014.

[36] Kyah Elisabeth Mercedes Smaal. Strategic manipulation in voting under higher-order reasoning. Master's thesis, Institute for Logic, Language and Computation, University of Amsterdam, 2019.

[37] Dale O Stahl and Paul W Wilson. Experimental evidence on players' models of other players. *Journal of economic behavior & organization*, 25(3):309–327, 1994.

[38] Hans van Ditmarsch, Jérôme Lang, and Abdallah Saffidine. Strategic voting and the logic of knowledge. In *11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 12)*, pages 1247–1248, 2012.

[39] Yuliya A Veselova. Does incomplete information reduce manipulability? *Group Decision and Negotiation*, 29(3):523–548, 2020.

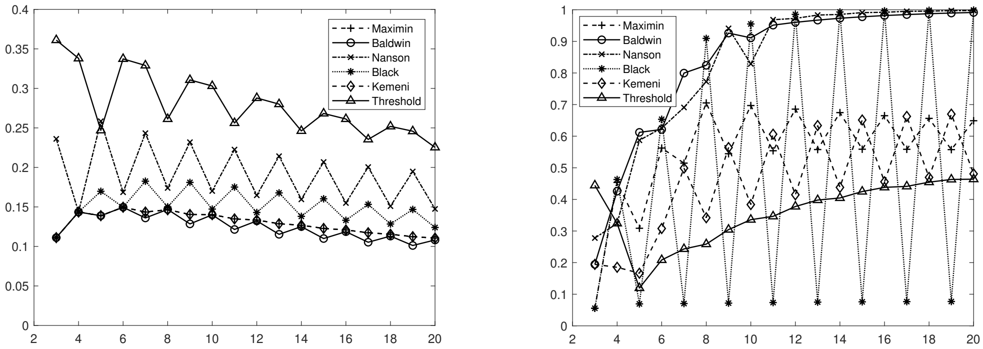
[40] Mark Wilson and R Reyhani Shokat Abad. The probability of safe manipulation. *COMSOC 2010*, 2010.

[41] Lirong Xia and Vincent Conitzer. Stackelberg voting games: Computational aspects and paradoxes. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 24, 2010.

Yuliya A. Veselova
National Research University Higher School of Economics
Moscow, Russia
V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences
Moscow, Russia
Email: `yul-r@mail.ru`

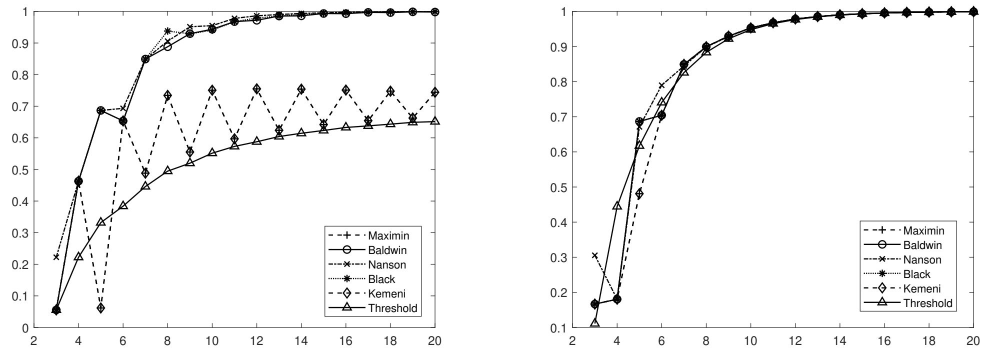
Daniel Karabekyan
National Research University Higher School of Economics
Moscow, Russia
Email: `dkarabekyan@hse.ru`

A Appendix



a) *Profile*-PIF b) *Rank*-PIF

Figure 3: The values of $I^{M1}(3, n, \pi, F)$



a) *Winner*-PIF b) *1Winner*-PIF

Figure 4: The values of $I^{M1}(3, n, \pi, F)$

B Appendix

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.06	.04	.03	.04	.03	.02	.03	.02	.02	.02	.02	.01	.01	.01	.01	.01	.01	.01
Veto	.06	.07	.08	.09	.09	.08	.08	.08	.07	.07	.07	.06	.06	.06	.05	.05	.05	.04
Borda	.00	.04	.03	.03	.02	.02	.01	.01	.01	.01	.01	.01	.01	.00	.00	.00	.00	.00
Run-off	.00	.00	.00	.00	.00	.01	.01	.00	.00	.00	.01	.00	.00	.00	.00	.00	.00	.00
STV	.06	.04	.00	.01	.01	.00	.03	.00	.00	.01	.00	.00	.01	.00	.00	.01	.00	.00
Copeland	.00	.04	.00	.03	.00	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
Maximin	.00	.00	.00	.00	.00	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
Baldwin	.00	.00	.00	.00	.00	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
Nanson	.00	.00	.07	.01	.06	.03	.04	.03	.03	.02	.02	.02	.01	.01	.01	.01	.01	.01
Black	.00	.00	.02	.00	.03	.01	.03	.01	.04	.01	.03	.01	.03	.01	.03	.01	.02	.01
Kemeny	.00	.00	.00	.00	.00	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
Threshold	.17	.04	.02	.12	.05	.05	.04	.05	.07	.04	.05	.06	.04	.04	.05	.04	.03	.04

Table 5: $I^{M1}(3,n,Profile,F) - I^{M2}(3,n,Profile,F)$

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.07	.12	.12	.18	.23	.25	.36	.31	.33	.43	.37	.39	.47	.40	.42	.50	.42
Veto	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda	.28	.19	.08	.08	.07	.06	.07	.08	.09	.10	.10	.11	.11	.12	.12	.12	.13	.13
Run-off	.11	.35	.28	.52	.29	.42	.36	.54	.27	.42	.24	.26	.10	.21	.00	.09	.00	.00
STV	.00	.07	.52	.24	.00	.72	.23	.36	.74	.37	.00	.58	.27	.00	.00	.21	.00	.00
Copeland	.19	.25	.21	.21	.25	.22	.27	.22	.29	.23	.29	.23	.30	.23	.30	.23	.30	.24
Maximin	.06	.21	.02	.16	.17	.15	.08	.09	.08	.00	.08	.00	.08	.00	.08	.00	.08	.00
Baldwin	.19	.19	.30	.32	.19	.40	.31	.27	.30	.32	.29	.33	.28	.34	.29	.32	.29	.32
Nanson	.28	.17	.23	.30	.00	.28	.29	.05	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Black	.06	.22	.02	.13	.03	.32	.04	.31	.03	.41	.04	.41	.04	.41	.04	.33	.04	.33
Kemeny	.19	.16	.15	.21	.37	.32	.42	.29	.45	.31	.48	.33	.49	.34	.50	.35	.51	.36
Threshold	.06	.17	.11	.19	.24	.23	.30	.34	.33	.37	.40	.39	.42	.44	.43	.45	.46	.45

Table 6: $I^{M2}(3,n,Rank,F)$

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.10	.12	.14	.26	.28	.25	.33	.33	.30	.37	.36	.34	.39	.37	.36	.40	.38
Veto	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda	.22	.31	.15	.15	.16	.17	.19	.20	.22	.23	.24	.25	.26	.27	.27	.28	.28	.29
Run-off	.22	.38	.71	.70	.65	.59	.55	.59	.25	.50	.35	.20	.30	.26	.00	.24	.00	.00
STV	.00	.10	.71	.29	.00	.60	.28	.14	.00	.38	.00	.00	.00	.00	.00	.00	.00	.00
Copeland	.06	.37	.00	.29	.00	.32	.00	.35	.00	.36	.00	.38	.00	.38	.00	.39	.00	.39
Maximin	.06	.22	.02	.13	.44	.12	.21	.09	.24	.00	.27	.00	.28	.00	.30	.00	.31	.00
Baldwin	.06	.22	.64	.13	.17	.42	.00	.09	.00	.08	.00	.08	.00	.08	.00	.07	.00	.07
Nanson	.22	.22	.25	.25	.00	.21	.04	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Black	.06	.22	.64	.13	.00	.32	.00	.45	.00	.17	.00	.16	.00	.15	.00	.07	.00	.07
Kemeny	.06	.22	.02	.13	.44	.12	.21	.09	.24	.00	.27	.00	.28	.00	.30	.00	.31	.00
Threshold	.06	.22	.32	.14	.19	.18	.23	.25	.24	.27	.29	.28	.30	.31	.30	.32	.32	.31

Table 7: $I^{M2}(3,n,Winner,F)$

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.15	.00	.17	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Veto	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda	.31	.27	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Run-off	.22	.52	.71	.71	.65	.32	.55	.45	.25	.26	.35	.00	.30	.00	.00	.00	.00	.00
STV	.00	.15	.71	.29	.00	.62	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Copeland	.17	.21	.00	.15	.00	.20	.00	.23	.00	.26	.00	.28	.00	.29	.00	.30	.00	.31
Maximin	.17	.18	.21	.57	.26	.63	.00	.24	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Baldwin	.17	.18	.43	.57	.17	.50	.00	.24	.00	.26	.00	.28	.00	.00	.00	.00	.00	.00
Nanson	.31	.18	.00	.57	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Black	.17	.18	.43	.57	.00	.63	.00	.24	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Kemeny	.17	.18	.21	.57	.26	.63	.00	.24	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Threshold	.11	.44	.62	.13	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 8: $I^{M2}(3,n,1Winner,F)$

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.00	.02	.03	.02	.03	.04	.03	.04	.04	.03	.04	.04	.03	.03	.03	.03	.03
Veto	.00	.01	.02	.03	.03	.05	.06	.07	.08	.09	.09	.10	.11	.11	.12	.12	.13	.13
Borda	.00	.00	.01	.03	.03	.05	.05	.06	.06	.07	.07	.07	.07	.07	.08	.08	.08	.08
Run-off	.00	.00	.00	.00	.01	.03	.02	.01	.01	.02	.02	.03	.02	.02	.02	.02	.03	.04
STV	.00	.00	.00	.00	.01	.00	.04	.07	.00	.02	.05	.02	.04	.07	.01	.03	.05	.03
Copeland	.00	.00	.00	.03	.00	.04	.00	.04	.00	.04	.01	.03	.01	.03	.01	.03	.01	.03
Maximin	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01	.01	.01	.01	.01	.01	.01	.01
Baldwin	.00	.00	.00	.00	.00	.00	.00	.00	.01	.00	.01	.01	.02	.01	.02	.01	.02	.02
Nanson	.00	.00	.01	.01	.08	.02	.12	.04	.14	.06	.16	.08	.16	.09	.17	.09	.17	.10
Black	.00	.00	.01	.00	.02	.00	.03	.01	.04	.01	.04	.02	.05	.02	.05	.03	.06	.03
Kemeny	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.01	.01	.01	.01	.01	.01	.01	.01
Threshold	.00	.02	.00	.06	.05	.03	.11	.07	.05	.11	.09	.08	.12	.11	.10	.13	.12	.12

Table 9: $I^{M2}(3,n,Profile,F) - I^{M3}(3,n,Profile,F)$

Rank	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.07	.12	.12	.18	.23	.25	.36	.31	.33	.43	.37	.39	.47	.40	.42	.50	.42
Veto	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda	.28	.19	.08	.08	.07	.06	.07	.08	.09	.10	.10	.11	.11	.12	.12	.12	.13	.13
Run-off	.11	.35	.28	.52	.29	.42	.36	.54	.27	.42	.24	.26	.10	.21	.00	.09	.00	.00
STV	.00	.07	.52	.24	.00	.72	.23	.22	.74	.37	.00	.58	.27	.00	.00	.21	.00	.00
Copeland	.19	.25	.21	.21	.25	.22	.27	.22	.29	.23	.29	.23	.30	.23	.30	.23	.30	.24
Maximin	.06	.21	.02	.16	.17	.11	.07	.00	.07	.00	.07	.00	.07	.00	.07	.00	.06	.00
Baldwin	.19	.19	.30	.32	.19	.28	.31	.18	.15	.23	.13	.22	.14	.23	.14	.24	.15	.25
Nanson	.28	.17	.19	.30	.00	.18	.25	.05	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Black	.06	.22	.02	.13	.02	.09	.03	.09	.03	.09	.03	.01	.04	.01	.04	.01	.04	.01
Kemeny	.19	.16	.15	.21	.36	.31	.42	.28	.45	.31	.48	.33	.49	.34	.50	.35	.51	.36
Threshold	.06	.17	.11	.19	.24	.22	.30	.34	.32	.37	.40	.38	.42	.44	.42	.45	.46	.45

Table 10: $I^{M3}(3,n,Rank,F)$

Winner	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.10	.12	.14	.26	.28	.25	.33	.33	.30	.37	.36	.34	.39	.37	.36	.40	.38
Veto	.03	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda	.22	.31	.15	.15	.16	.17	.19	.20	.22	.23	.24	.25	.26	.27	.27	.28	.28	.29
Run-off	.22	.38	.71	.70	.37	.59	.00	.37	.00	.08	.00	.00	.00	.00	.00	.00	.00	.00
STV	.00	.10	.71	.29	.00	.60	.00	.00	.00	.15	.00	.00	.00	.00	.00	.00	.00	.00
Copeland	.06	.37	.00	.29	.00	.32	.00	.35	.00	.36	.00	.38	.00	.38	.00	.39	.00	.39
Maximin	.06	.22	.02	.13	.43	.09	.20	.00	.23	.00	.25	.00	.27	.00	.28	.00	.29	.00
Baldwin	.06	.22	.64	.13	.00	.30	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Nanson	.22	.22	.21	.25	.00	.18	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Black	.06	.22	.64	.13	.00	.09	.00	.09	.00	.08	.00	.00	.00	.00	.00	.00	.00	.00
Kemeny	.06	.22	.02	.13	.43	.09	.20	.00	.23	.00	.25	.00	.27	.00	.28	.00	.29	.00
Threshold	.06	.22	.32	.14	.19	.18	.22	.25	.23	.27	.29	.27	.30	.31	.29	.32	.32	.31

Table 11: $I^{M3}(3,n,Winner,F)$

	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Plurality	.00	.15	.00	.17	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Veto	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Borda	.31	.27	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Run-off	.22	.52	.71	.57	.37	.32	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
STV	.00	.15	.71	.29	.00	.22	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Copeland	.17	.21	.00	.15	.00	.20	.00	.23	.00	.26	.00	.28	.00	.29	.00	.30	.00	.31
Maximin	.17	.18	.21	.57	.26	.20	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Baldwin	.17	.18	.43	.57	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Nanson	.31	.18	.00	.57	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Black	.17	.18	.43	.57	.00	.20	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Kemeny	.17	.18	.21	.57	.26	.20	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00
Threshold	.11	.44	.62	.13	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00	.00

Table 12: $I^{M3}(3,n,1Winner,F)$

C Appendix

Proposition C.1. *For any PIF π, for any rule F, the number of voters n, and the number of alternatives m if $I^{M1}(m,n,F,\pi')=0$, then $I^{M2}(m,n,\pi',F)=0$.*

Proof. The proof follows directly from the definitions of manipulation in Model 1 and Model 2. To have an incentive to manipulate in Model 2, a voter needs to have an incentive to manipulate in Model 1. If $I^{M1}(m,n,\pi',F)=0$, then no voter has in incentive to manipulate in Model 1, and, consequently, also not in Model 2. So, $I^{M2}(m,n,\pi',F)=0$. □

Lemma C.1. *For any scoring rule F, any number of alternatives m, any voter i, any jump Δ_j, and any two distinct alternatives a_h and a_l there is a number of voters n^*, such that for all $n>n^*$ there exists $\boldsymbol{P}\in L(X)^N$ such that*
1) $P_i=(a_1,a_2,...,a_m)$;
2) for all $a_g\in X\setminus\{a_h,a_l\}$ $S(a_h,\boldsymbol{P})>S(a_g,\boldsymbol{P})$;
3) $S(a_h,\boldsymbol{P})-S(a_l,\boldsymbol{P})=\Delta_j$ ($S(a_h,\boldsymbol{P})-S(a_l,\boldsymbol{P})=0$).

Proof. 1) For simplicity, let voter i be the first with a preference order $P^1=(a_1,...,a_h,...,a_l,...,a_m)$ (which means $a_1P^1a_hP^1a_lP^1a_m$, dots mean there can be other alternatives). First, we consider the case when $h<l$. Then we denote $P^2=(a_l,...,a_1,...,a_h,...)$, $P^3=(a_h,...,a_l,...,a_1,...)$ (alternatives a_1, a_h, and a_l are on the same places, but cycled according to a permutation $(a_1\,a_l\,a_h)$), $P^4=(a_1,...,a_l,...,a_h,...,a_m)$ (the same as P^1, but a_h and a_l switched), $P^5=(a_h,a_l,...)$, $P^6=(a_l,a_h,...)$, $P^7=(...,a_h|a_l,...)$ (the line $|$ denotes the position of j-th jump, Δ_j).

2) Let us prove by construction that there exists a profile with $S(a_h,\mathbf{P})-S(a_l,\mathbf{P})=\Delta_j$. For an odd n: $\mathbf{P}'=(P^1,P^4,qP^5,qP^6,P^7)$. For an even n: $\mathbf{P}''=(P^1,P^2,P^3,qP^5,qP^6,P^7)$. A profile with $S(a_h,\mathbf{P})=S(a_l,\mathbf{P})$ is constructed the same way by leaning out P^7.

3) Now we prove that the condition $\forall a_g\in X\setminus\{a_h,a_l\}$ $S(a_h,\mathbf{P})>S(a_g,\mathbf{P})$ could be satisfied for the constructed profiles. First, in preferences of type P^5 and P^6 let all other $m-2$ alternatives be cycled. The number of scores got by a_h and a_l in (qP^5,qP^6) is qs_1+qs_2. Let $h=[2q/(m-2)]$, which is the number of whole cycles in (qP^5,qP^6). The number of scores got by any alternative from $X\setminus\{a_h,a_l\}$ is not greater than $h(s_3+...+s_m)+(2q-h(m-2))s_3\le hs_m+(2q-h)s_3$. Since $s_1\ge s_2\ge s_3\ge...\ge s_m$ and $s_1>s_m$, $qs_1+qs_2>hs_m+(2q-h)s_3$ and the difference is not less then $min(h,q)(s_1-s_m)$ (so, for this difference to be positive for all alternatives in $X\setminus\{a_h,a_l\}$, there should be at least one cycle of $m-2$ alternatives, i.e. $m-2<2q$). Thus, by taking q big enough we can make scores of a_h in $\mathbf{P}$ be higher than scores of any other alternative in each type of $\mathbf{P}$ constructed in 2). If the condition $\forall a_g\in X\setminus\{a_h,a_l\}$ $S(a_h,\mathbf{P})>S(a_g,\mathbf{P})$ is satisfied for some q, then for $q'=q+1$ it is also satisfied.

$$n'=3+min\{q:\forall a_g\in X\setminus\{a_h,a_l\}S(a_h,\mathbf{P}')>S(a_j,\mathbf{P}')\},$$

$$n''=4+min\{q:\forall a_g\in X\setminus\{a_h,a_l\}S(a_h,\mathbf{P}'')>S(a_g,\mathbf{P}'')\},$$

$$n^*=max(n',n'').$$

Therefore, for all $n>n^*$ there exists a preference profile with $P_i=(a_1,a_2,...,a_m)$, $\forall a_g\in X\setminus\{a_h,a_l\}$ $S(a_h,\mathbf{P})>S(a_g,\mathbf{P})$ and $S(a_h,\mathbf{P})-S(a_l,\mathbf{P})=\Delta_j$.

4) If $h>l$, i.e. a_h is less preferred than a_l by voter i, then we switch alternatives a_h and a_l in P^1, P^2, P^3, and P^4 with all other parts of the proof staying the same. □

Theorem C.1. *For any scoring rule F and any number of alternatives m there is a finite number of voters n^*, such that for all $n>n^*$ it holds $I^{M2}(m,n,1Winner,F)=0$.*

Proof. Let $X = \{a_1,...,a_m\}$. Consider a scoring rule with a scoring vector $s = (s_1,s_2,...,s_m)$, the first jump in s goes after s_k, $s_k - s_{k+1} = \Delta_1$.

1) Let us prove that voter i with preferences $a_1P_ia_2P_i...P_ia_m$ has no incentive to manipulate (in Model 1) under *1Winner*-PIF if $V(\mathbf{P}) \in \{a_1,a_2,...,a_{k+1}\}$.

1.1) If $V(\mathbf{P}) = a_1$, then there is no need for voter i to misrepresent preferences, since it is the best alternative for i.

1.2) Suppose that $V(\mathbf{P}) = b$, $b \in \{a_2,a_3,...,a_{k+1}\}$ and i manipulates in favor of some a, such that aPb. If i puts alternative a higher (if a is not a_1), then nothing changes for a since $s_1 = ... = s_k$. Thus, i could only put b lower in $\tilde{P}_i$, but then some alternative $c \in \{a_{k+2},...,a_m\}$ goes higher. If $b = a_{k+1}$ and there are no jumps in s after $k+1$, then putting b lower will not change the scores of b and c and no manipulation is possible in this case. In other cases b gets $-A$ scores and c gets $+A$ scores, where $A = \alpha_1\Delta_1 + \alpha_2\Delta_2 + ...$ and $\alpha_j \in \{0,1\}$.

Thus, if there exists $\mathbf{P}' = (P_i,\mathbf{P}'_{-i})$, $\mathbf{P}'_{-i} \in W_i^{\pi(\mathbf{P})}$, such that $S(b,\mathbf{P}') - S(c,\mathbf{P}') = min(\Delta_1,\Delta_2)$, then c wins in $(\tilde{P}_i,\mathbf{P}'_{-i})$, since it gets $+A$ scores and b gets $-A$ scores. By 1 such preference profile exists for all $n > \hat{n}$. It means that there is a chance of getting c as a result which is worse than b for i. Therefore, i does not have an incentive to manipulate in Model 1 under *1Winner*-PIF when $V(\mathbf{P}) \in \{a_1,a_2,...,a_{k+1}\}$.

2) Now prove that if $V(\mathbf{P}) \in \{a_{k+2},...,a_m\}$, then voter i with preferences $a_1P_ia_2P_i...P_ia_m$ has an incentive to manipulate (in Model 1) under *1Winner*-PIF. Suppose, $V(\mathbf{P}) = c$ and c is on $k+t$-th place in P_i, where $t \in \{2,...,m\}$. Then manipulation in favor of some $b \in \{a_{k+1},...,a_{k+t-1}\}$ is possible: voter i switches alternatives $a \in \{a_1,...,a_k\}$ and b. Since it does not matter which $a \in \{a_1,...,a_k\}$ to choose for switching with b, we can assume that it is a_k. By the principle of the best winning alternative b must be a_{k+1}. After this manipulation b gets $+\Delta_1$ scores and a gets $-\Delta_1$ scores, scores of other alternatives do not change. Thus, if there exists $\mathbf{P}' = (P_i,\mathbf{P}'_{-i})$, $\mathbf{P}'_{-i} \in W_i^{\pi(\mathbf{P})}$, such that $S(c,\mathbf{P}') - S(b,\mathbf{P}') = \Delta_1$, then b wins in $(\tilde{P}_i,\mathbf{P}'_{-i})$ provided that bP_Tc. If cP_Tb, then we need a profile $\mathbf{P}'' = (P_i,\mathbf{P}''_{-i})$, $\mathbf{P}''_{-i} \in W_i^{\pi(\mathbf{P})}$, such that $S(c,\mathbf{P}'') = S(b,\mathbf{P}'')$ for b to win after getting $+\Delta_1$ scores. By 1, such profiles exist for all $n > \check{n}$. Therefore, in some preference profiles of i's information set b wins, and there is no risk of getting a worse alternative as a result, so, i has an incentive to manipulate in Model 1 under *1Winner*-PIF when $V(\mathbf{P}) \in \{a_{k+2},...,a_m\}$ for all $n > \check{n}$.

3) Take a voter i with preferences $a_1P_ia_2P_i...P_ia_m$ having an incentive to manipulate (in Model 1). If in $W_i^{\pi(\mathbf{P})}$ there is at least one preference profile such that $V(P_i,\tilde{\mathbf{P}}'_{-i})P_iV(\tilde{P}_i,\tilde{\mathbf{P}}'_{-i})$, then voter i has no incentive to manipulate in Model 2. We prove the existence of such a profile $\mathbf{P}'$ by construction.

For simplicity, let us denote $c = a_{k+t}$, $b = a_{k+1}$, $a = a_k$.

We construct $\mathbf{P}'$ with the following preference orders: preferences of voter i $P^1 = (...a|b...c...)$, $P^2 = (...b|a...c...)$, $P^3 = (...|c...a...)$, $P^4 = (...c|...a...)$ (in P^1 and P^2 c on $k+t$-th place, in P^3 and P^4 a on $k+t$-th place), $P^5 = (c\,a...|...)$, $P^6 = (a\,c...|...)$, $P^7 = (c|a...)$. Thus, we have voter i with a preference order P^1, having an incentive to manipulate in Model 1. Voters with preference order P^2 are also GS-manipulators, but their manipulation is an opposite one (putting b on $k+1$-th place and a on k-th). Voters with preference orders P^3, P^4, P^5, and P^6 do not have an incentive to manipulate in Model 1 since the winning alternative is not lower than $k+1$ place.

The way of construction depends on a tie-breaking between c and a. Consider two cases: aP_Tc and cP_Ta. For all the following cases let us assume that the condition $\forall x \in X \setminus \{a,c\}$ $S(c) > S(x)$ is satisfied. Later (in part 6 of the proof) we will show that there is a finite number of voters n^*, such that for all $n > n^*$ this is true.

4) If aP_Tc, then $S(c) > S(a)$.
4.1) n is even, $k \in \{1,\dots,m-2\}$, $\mathbf{P}' = (P^1, 2P^2, 2P^3, P^4, qP^5, qP^6, 2P^7)$. For this profile, $S(c,\mathbf{P}') - S(a,\mathbf{P}') = 2\Delta_1$ and $V(\mathbf{P}') = c$. If all voters from $\Pi(\mathbf{P}')$ manipulate, then $S(c,\tilde{\mathbf{P}}') - S(a,\tilde{\mathbf{P}}') = \Delta_1$. If i does not manipulate, and $\Pi(\mathbf{P}') \setminus \{i\}$ does, then $S(c,(P'_i,\tilde{\mathbf{P}}'_{-i})) = S(a,(P'_i,\tilde{\mathbf{P}}'_{-i}))$ and $V(P'_i,\tilde{\mathbf{P}}'_{-i}) = a$.
4.2) n is odd, $k \in \{1,\dots,m-2\}$, $\mathbf{P}' = (P^1, P^2, P^3, P^4, qP^5, qP^6, P^7)$. For this profile, $S(c,\mathbf{P}') - S(a,\mathbf{P}') = \Delta_1$ and $V(\mathbf{P}') = c$. If all voters from $\Pi(\mathbf{P}')$ manipulate, then $S(c,\tilde{\mathbf{P}}') - S(a,\tilde{\mathbf{P}}') = \Delta_1$. If i does not manipulate, and $\Pi(\mathbf{P}') \setminus \{i\}$ does, then $S(c,(P'_i,\tilde{\mathbf{P}}'_{-i})) = S(a,(P'_i,\tilde{\mathbf{P}}'_{-i}))$ and $V(P'_i,\tilde{\mathbf{P}}'_{-i}) = a$.
5) If cP_Ta, then $S(a) \geq S(c)$.
5.1) n is even, $k \in \{1,\dots,m-2\}$, $\mathbf{P}' = (P^1, P^2, P^3, P^4, qP^5, qP^6)$. For this profile, $S(a,\mathbf{P}') = S(c,\mathbf{P}')$ and $V(\mathbf{P}') = c$. If all voters from $\Pi(\mathbf{P}')$ manipulate, then $S(c,\tilde{\mathbf{P}}') = S(a,\tilde{\mathbf{P}}')$ and again $V(\tilde{\mathbf{P}}') = c$. If i does not manipulate and $\Pi(\mathbf{P}') \setminus \{i\}$ (which is only one voter with preferences P^2 in this case) does, then $S(c,(P'_i,\tilde{\mathbf{P}}'_{-i})) < S(a,(P'_i,\tilde{\mathbf{P}}'_{-i})) = S(a,\mathbf{P}') + \Delta_1$ and $V(P'_i,\tilde{\mathbf{P}}'_{-i}) = a$.
5.2) n is odd, $k = 1$, $\mathbf{P}' = (P^1, 2P^2, 2P^3, P^4, (q+1), qP^6)$. For this profile, $S(c,\mathbf{P}') - S(a,\mathbf{P}') = \Delta_1$ and $V(\mathbf{P}') = c$. If all voters from $\Pi(\mathbf{P}')$ manipulate, then $S(a,\tilde{\mathbf{P}}') = S(a,\mathbf{P}') + 2\Delta_1 - \Delta_1$, $S(c,\tilde{\mathbf{P}}') = S(a,\tilde{\mathbf{P}}')$ and again $V(\tilde{\mathbf{P}}') = c$. If i does not manipulate, $S(c,(P'_i,\tilde{\mathbf{P}}'_{-i})) < S(a,(P'_i,\tilde{\mathbf{P}}'_{-i})) = S(a,\mathbf{P}') + 2\Delta_1$ and $V(P'_i,\tilde{\mathbf{P}}'_{-i}) = a$.
5.3) n is odd, $k \in \{2,\dots,m-2\}$, $\mathbf{P}' = (P^1, P^2, P^3, P^4, (q+1)P^5, qP^6)$. All relations between scores of c and a are the same as in 5.1).

6) Now we need to show that it is possible to have the condition $\forall x \in X \setminus \{a,c\}\ S(c) > S(x)$ satisfied and find out when it is possible. Let us assume that in preferences of type P^5 and P^6 all other $m-2$ alternatives are cycled, i.e. $(a\,c\,x_1\,x_2\,...x_{m-1})$, $(a\,c\,x_{m-1}\,x_1\,x_2...x_{m-2})$, etc. Let $h = [2q/(m-2)]$, which is the number of whole cycles in (qP^5, qP^6).

The number of scores got by c in (qP^5, qP^6) is $qs_1 + qs_2$. The number of scores got by any alternative from $X \setminus \{a,c\}$ is not greater than $h(s_3 + ... + s_m) + (2q - h(m-2))s_3 \leq hs_m + (2q-h)s_3$. Since $s_1 \geq s_2 \geq s_3 \geq ... \geq s_m$ and $s_1 > s_m$, $qs_1 + qs_2 > hs_m + (2q-h)s_3$ and the difference is not less then $min(h,q)(s_1 - s_m)$. Thus, by taking q big enough we can make scores of c in $\mathbf{P}$ be higher than scores of any other alternative.

Let such number of voters that the condition $\forall x \in X \setminus \{a,c\}\ S(c) > S(x)$ is satisfied for all cases 4.1), 4.2), 5.1), 5.2) be denoted by $\grave{n}$. Summing up, having a fixed number of alternatives, voter i with preferences $a_1P_ia_2P_i...P_ia_m$ has no incentive to manipulate (in Model 1) under a scoring rule and *1Winner*-PIF if $V(\mathbf{P}) \in \{a_1, a_2, ..., a_{k+1}\}$ for all $n > \hat{n}$, but has an incentive to manipulate if $V(\mathbf{P}) \in \{a_{k+2}, ..., a_m\}$ for all $n > \check{n}$. However, voter i does not have an incentive to manipulate under the same conditions in Model 2 for all $n > \grave{n}$. Thus, $I^{M2}(m,n,1Winner,F) = 0$ for all $n > n^* = max(\hat{n},\check{n},\grave{n})$. □

Proposition C.2. *For any PIF π, for any rule F, the number of voters n, and the number of alternatives m if $I^{M1}(m,n,\pi,F) = 0$, then $I^{M3}(m,n,\pi,F) = 0$.*

Proof. The proof is the same as for Proposition 1, but with respect to Model 3. □

Proposition C.3. *For any PIF π, for any rule F, the number of voters n, and the number of alternatives m if $I^{M2}(m,n,\pi,F) = 0$, then $I^{M3}(m,n,\pi,F) = 0$.*

Proof. If $I^{M2}(m,n,\pi,F) = 0$, then in any preference profile $\mathbf{P}$ no voter i has an incentive to manipulate in Model 2. There are two cases. In the first case voter i does not have an incentive to manipulate in Model 1. Then, by Definition 4, she does not have an incentive to manipulate in Model 3. In the second case, voter i has an incentive to manipulate in Model 1, but there is $\mathbf{P}'_{-i}$ in W_i^{π} such that

$V(P_i,\tilde{\mathbf{P}}'_{-i})P_iV(\tilde{P}_i,\tilde{\mathbf{P}}'_{-i})$ for all $\tilde{P}_i$. Thus, condition ii) from Definition 4 is not satisfied, and voter i does not have an incentive to manipulate in Model 3. Consequently, $I^{M3}(m,n,\pi,F)=0$. $\square$

Proposition C.4. *Suppose, π' is at least as informative as π''. Then for any $M \in \{M1,M2,M3\}$, for any rule F, number of voters n, and number of alternatives m if $I^M(m,n,\pi',F)=0$, then $I^M(m,n,\pi'',F)=0$.*

Proof. Let us introduce a notation specially for this proof: every voter has her own information set W_i^{π}, but in accordance with a behavioral model M voters may change their preferences (as in Model 2 and 3) or not (Model 1), so let $\hat{W}_i^{\pi}$ be an information set of voter i with respect to the behavioral model M.

If $I^M(m,n,F,\pi')=0$, then for any voter i in every profile $\mathbf{P}$ and any strategy $\tilde{P}_i \neq P_i$ there exists $\mathbf{P}'_{-i}$ in $\hat{W}_i^{\pi}$, such that $V(P_i,\mathbf{P}'_{-i})P_iV(\tilde{P}_i,\mathbf{P}'_{-i})$. Since π' is at least as informative as π'', $W_i^{\pi'} \subseteq W_i^{\pi''}$, and, consequently, $\hat{W}_i^{\pi'} \subseteq \hat{W}_i^{\pi''}$. Then, $\mathbf{P}'_{-i}$ belongs to $\hat{W}_i^{\pi''}$ as well and voter i does not have an incentive to manipulate under PIF π''. Since that holds for any voter i in any preference profile $\mathbf{P}$ and any strategy $\tilde{P}_i \neq P_i$, we get $I^M(m,n,\pi'',F)=0$. $\square$