Beyond Gibbard and Satterthwaite: Voting Manipulation Games

Arkadii Slinko

Abstract The Gibbard–Satterthwaite theorem implies the existence of voters, called manipulators, who can change the election outcome in their favour by voting strategically. However, when a given preference profile admits several such manipulators, voting becomes a game played by these voters. They have to reason strategically about each other's actions.

Voting is a common method of collective decision making, which enables the participating voters to identify the best candidate for the society given their individual rankings of the candidates. However, as early as Farquharson (1969), it was noticed that for most common rules voters sometimes can misrepresent their preference and improve the outcome for themselves. In social choice this is now called a manipulation, in political science this is called tactical or strategic voting. Pattanaik (1973) conjectured that no "reasonable" voting rule is immune to manipulation. This indeed was shown independently by Gibbard (1973) and Satterthwaite (1975): if there are at least 3 candidates, then any onto, non-dictatorial voting rule admits a preference profile (a collection of voters' rankings) where some voter would be better off by submitting a ranking that differs from his truthful one. In other words, the sincere profile of preferences is not a Nash equilibrium.

The problem may be further exacerbated by the presence of a number of such voters —we will call them Gibbard–Satterthwaite manipulators, or GS-manipulators. Indeed, if several such voters—attempt to manipulate the election simultaneously in an uncoordinated fashion, the outcome may differ not just from the outcome of the truthful voting, but also from the outcome that any of the GS-manipulators was trying to achieve, due to complex interference among the different manipulative votes. The outcome may be undesirable for all manipulators. In other word, if voters are strategic, voting becomes a simultaneous one-shot game. Here is one basic example.

Arkadii Slinko was supported by the Marsden Fund 3706352 of The Royal Society of New Zealand.

A. Slinko (✉)
The University of Auckland, Auckland, New Zealand
e-mail: a.slinko@auckland.ac.nz

J.-F. Laslier et al. (eds.), *The Future of Economic Design*, Studies in Economic Design, https://doi.org/10.1007/978-3-030-18050-8_19

Example 1 Suppose four people are to choose between three alternatives. Let the profile of sincere preferences be

1	2	3	4
a	b	c	c
b	a	a	b
c	c	b	a

and the rule used be Plurality with breaking ties in accord with the order $a > b > c$. If everybody votes sincerely, then c is elected. Voters 1 and 2 are the Gibbard–Satterthwaite manipulators at this profile since voters 3 and 4 get their best possible outcome and are expected to vote sincerely. Voter 1 can make b to win by voting $b > a > c$ and voter 2 can make a to win by voting $a > b > c$. However, if they both try to manipulate, c will remain the winner. Each of them would prefer that the other one manipulates. If, for example, for voters 1 and 2 the utility of their top preference is 2, the utility of their middle preference is 1 and the utility of their bottom preference is 0 (these are ordinal, not cardinal), these voters are playing the game which normal form is

	s_2	i_2
s_1	0, 0	2, 1
i_1	1, 2	0, 0

where voter 1 is the row player and voter 2 is the column player, s_1, s_2 are their sincere votes and i_1, i_2 are their manipulative votes.

There is a substantial body of research dating back to Farquharson (1969) that explores the consequences of modeling non-truthful voting as a strategic game; see, also Fishburn (1978), Moulin (1979), Feddersen et al. (1990), Myerson and Weber (1993), De Sinopoli (2000), Dhillon and Lockwood (2004), Sertel and Sanver (2004), De Sinopoli et al. (2015), Desmedt and Elkind (2010), Obraztsova et al. (2013) to mention a few. The main common feature of them is that the utility of a voter depends solely on the winner of the election. The features of the existing models that vary are:

- voters' utilities maybe ordinal or be von-Neumann-Morgenstern utilities;
- voters may use randomised strategies or not;
- voters may be fully or only boundedly rational.

The most popular frameworks so far has been the one investigated by Moulin (1981) with ordinal utilities and Myerson and Weber (1993) with von-Neumann-Morgenstern utilities. The latter model, in particular, stipulate that each voter has a utility for the election of each candidate. Both Moulin and Myerson and Weber suggested the use of Nash equilibrium as a solution concept for the analysis of voting games, however, acknowledging that sometimes this idea led to a large number of Nash equilibria.

The use of Nash equilibrium in analysis of voting games was widely criticised (see, e.g., De Sinopoli 2000) and the classical example that wanders about from one paper supporting this criticism to another is as follows.

Example 2 Suppose each of $n \geq 3$ voters has the same preference order. Then voting for the least preferred alternative for each of them is a Nash equilibrium.

In further works many attempts have been made to weed weird equilibria out. Farquharson (1969) was aware of the problem and suggested the sophisticated voting principle: reasonable equilibria must survive iterative deletion of dominated strategies. The following methods were considered: equilibria refinements (De Sinopoli 2000), costly voting (Sinopoli and Iannantuoni 2005), truth-biased voters Obraztsova et al. (2013), generic utilities (De Sinopoli 2001; De Sinopoli et al. 2015). The problem however remains not completely solved. These attempts, to my mind, are futile. What actually Example 2 shows is that it is not the solution concept that is to blame but the implicit assumption that voters may behave irrationally. In reality, however, voters may not be fully rational but they are certainly not irrational. And it is not a coincidence that in all aforementioned papers voters' reasoning about the voting situation and other voters' possible actions was absent.

The outlined approach always assumes that voters can submit any linear order as their vote (no matter how rational it is) without any regard to appraisal. Moulin (1981) divides voters into prudent and sophisticated depending on the amount of information about other players that they have. Effectively, if you have no information about other players you vote prudently, e.g., sincere. However, if you have all the information you need for sophisticated voting but you are unable to process it, you are back to the situation of no information.

Apart from the practical difficulty of sophisticated voting, political science has observed that some voters may be *ideological* and interested in stating their preference no matter how much information they have. Some observations tell us that the number of ideological voters is significant. In the famous Florida 2000 vote, when Bush won over Gore by just 537 votes 97,488 Nader supporters voted for Nader while in such a close election every strategic voter would vote either Gore or Bush and the overwhelming majority of Nader supporters preferred Gore to Bush.

Hence another approach to analysis of voting games would be not to assume that voters miraculously always end up in one of the Nash equilibria but to take their reasoning about other players as a starting point of the analysis. Naturally we have to assume that voters are only boundedly rational and their strategies depend on the level of depth of their reasoning. These voters need a good reason to abandon their sincere preference in favour of an insincere one.

In many papers—especially those discussing indices of manipulability (see, e.g., often cited Aleskerov and Kurbanov 1999)—it is implicitly assumed that voters manipulate as soon as they discover that they are in the position to manipulate provided everyone else will vote sincerely, which means that no reasoning about actions of others is assumed. Slinko and White (2008, 2014) were the first to assume that voters make a small step in reasoning about others, namely, about other voters who have exactly the same preference order (and assuming all the rest vote sincerely). According to Slinko and White such voters reason: "I am thinking about manipulating but other voters of my type[1] must think about this too as they are in the same position as me.[2] I have to interact with them." This is illustrated in the following example.

Example 3 Suppose four people are to choose between three alternatives. Let the profile of sincere preferences be

1	2	3	4
a	a	b	c
b	b	c	b
c	c	a	a

and the rule to be used is Borda with breaking ties in accord with the order $a > b > c$. If everybody votes sincerely, then b is elected. Voters 1 and 2 are Gibbard–Satterthwaite manipulators. Voter 1 can make a to win by voting $a > c > b$ and voter 2 can do the same. However, if they both try to manipulate, their worst alternative c will become the winner. Thus, these voters are playing the game which normal form is

	s_2	i_2
s_1	1, 1	2, 2
i_1	2, 2	0, 0

which is a sort of an anti-coordination game.

We note that in Example 3 it is easier for the two would-be manipulators to agree on the course of action than in Example 1 since they have the same type, i.e., have identical sincere preferences, and hence no conflict of interest but the absence of communication devices prevents them from doing this. Most mis-coordinations in our framework can be classified as instances of either strategic overshooting (too many voted strategically) or strategic undershooting (too few). If mis-coordination can result in strategic voters ending up worse off than they would have been had they all just voted sincerely, we call the strategic vote *unsafe*. Slinko and White (2008, 2014) showed that under every onto and non-dictatorial social choice rule there exist circumstances where a voter has an incentive to cast a safe strategic vote. Thus they extended the Gibbard–Satterthwaite Theorem by proving that every onto and non-dictatorial social choice rule can be individually manipulated by a voter casting a safe strategic vote.

Example 3 illustrates an important point. The game has two Nash equilibria but there is no way voters 1 and 2 can derive from their votes which Nash equilibria to play. Myerson and Weber (1993) claim that opinion polls can serve as coordination device for voters which is partly true; however sometimes an aggregated information of opinion polls may be insufficient.

[1]I.e., voters with the same order of appreciation of candidates.

[2]Assuming anonymity.

The generalisation of the Gibbard–Satterthwaite theorem given by Slinko and White prompts us to reconsider the concept of a manipulable profile and all the business related to indices of manipulability. For example, one may ask: Are the profiles in Examples 1 and 3 really manipulable?

Assuming that all voters of types different from their own are sincere (which is most likely untrue), the voter in the model of Slinko and White is (strongly) boundedly rational and has a much simplified view of the game. When we deal with boundedly rational players we have to have it in mind that their view of the game may be different from the game itself. The more rationality we assume, the more voter's view of the game is closer to reality. The next level of rationality of the voter would be for him to consider *all* GS-manipulators (not only of his own type) as players. At the previous level of rationality voters 1 and 2 in Example 1 would consider each other as sincere but at this next level they realise that there is a game to play.

The game played by GS-manipulators was studied in Elkind et al. (2015), Grandi et al. (2017). This means voters may strategise only when they are Gibbard–Satterthwaite manipulators and they can identify all other manipulators and try to optimise their vote relative to the information about those manipulators and their potential manipulations. These games for Plurality voting rule appear to be relatively simple and, in particular, always have a Nash equilibrium. Grandi et al. (2017) identify natural conditions implying the existence of Nash equilibria for k-approval with $k = 2, 3$. It appeared that some additional mild rationality conditions are necessary for 2-approval voting manipulation game to have a Nash equilibria. If voters are erratic and use unsound strategies, then a Nash equilibria may not exist. For 3-approval the sufficient conditions are stronger (but it is not clear if they are necessary). However, Elkind et al. (2015), Grandi et al. (2017) showed that even the so-called minimality conditions (which require that voters resort only to minimal manipulations), fail to ensure the existence of Nash equilibria for 4-approval voting rule.

One way to increase sophistication of voters is to allow countermanipulations. This means that a voter who cannot manipulate himself can vote insincerely in order to mitigate the damage that can be done by a GS-manipulator when he manipulates. The games with participation of a countermanipulator by their nature usually do not have any Nash equilibria.

Example 4 In a 2-by-2 game with one manipulator and one countermanipulator we do not necessarily have a Nash equilibrium. Suppose that the voting rule is 2-Approval and the profile is $V = (adcb, bdca)$ with ties broken according to $a > b > c > d$. For voter 1 switching c and d is a manipulation in favour of a. Voter 2 cannot manipulate but can countermanipulate switching c and d making c to win in case voter 1 manipulates. But then voter 1 would be better off switching c and d back after which the same move will be beneficial for voter 2. The normal form for such game would be

	s_2	i_2
s_1	2, 2	3, 0
i_1	3, 0	1, 1

We note that any 2-by-2 game, where two manipulators play, always has a Nash equilibrium, thus a manipulator-countermanipulator 2-by-2 games are very different.

Elkind et al. (2017) modeled bounded rationality of voters differently. They also assume that the voters reason about potential actions of other voters but for modeling boundedly rational they use an adaptation of the cognitive hierarchy model. They take non-strategic (sincere) voters as those belonging to level 0. The players of level 1 give best response assuming that all other players belong to level 0, and, when this particular voter is not a Gibbard–Satthethwaite manipulator it is defined to be the sincere vote. The players of level 2 give their best responce to assuming that all other players belong to level 0 or level 1. We note that players of level 2 are already quite sophisticated. They can, for example, think of countermanipulating or they can strategically stay sincere when they can manipulate. The emphasis of the paper by Elkind et al. (2017) is on the complexity of a level 2 voter deciding whether his manipulative strategy weakly dominates his sincere strategy. They present a polynomial time algorithm for 2-approval but prove NP-hardness for 4-Approval voting rule. The case of 3-approval remains open.

The algorithmic aspects of voting games with fully rational voters have also recently received some attention (Desmedt and Elkind 2010; Xia and Conitzer 2010; Thompson et al. 2013; Obraztsova et al. 2013). Empirical analysis of Nash equilibria in plurality election has been done in Thompson et al. (2013).

Conclusion

The study of voting manipulation games is in its infancy and it is a long way before we can get any realistic models of elections. It is clear that the assumption that voters are fully rational and that every election end up in one of the nice Nash equilibria (or one of the similar solution concepts) is not realistic. However little is known about what happens in reality. For example, it is extremely hard to estimate how many strategic voters are there in any election but the percentage of those who actually manipulated is easier to estimate; for example, Kawai and Watanabe (2013) estimate the number of such voters[3] in Japanese elections as between 2.5 and 5.5%. Moreover, Benjamin et al. (2013), show that preference misrepresentation is related to cognitive skills, and Choi et al. (2014) demonstrate that decision-making ability in laboratory experiments correlates strongly with socio-economic status and wealth. So it may be reasonable to assume that only a small fraction of voters in any election is strategic. Circumstantial evidence exists that ideological voters are always present in non-negligible numbers but there are no experimental data in this respect.

We need to further develop models of bounded rationality of voters. But primarily now we need more experimental work for better understanding of voters and their behaviour.

[3]They call such voters misaligned.

References

Aleskerov, F., & Kurbanov, E. (1999). Degree of manipulability of social choice procedures. In *Current trends in economics* (pp. 13–27). Berlin: Springer.

Benjamin, D. J., Brown, S. A., & Shapiro, J. M. (2013). Who is behavioral? Cognitive ability and anomalous preferences. *Journal of the European Economic Association*, *11*, 1231–1255.

Choi, S., Kariv, S., Muller, W., & Silverman, D. (2014). Who is (more) rational? *American Economic Review*, *104*, 1518–1550.

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

De Sinopoli, F. (2001). On the generic finiteness of equilibrium outcomes in plurality games. *Games and Economic Behavior*, *34*(2), 270–286.

De Sinopoli, F., Iannantuoni, G., & Pimienta, C. (2015). On stable outcomes of approval, plurality, and negative plurality games. *Social Choice and Welfare*, *44*(4), 889–909.

Desmedt, Y., & Elkind, E. (2010). Equilibria of plurality voting with abstentions. In *Proceedings of the 11th ACM conference on Electronic Commerce (EC-2010)* (pp. 347–356).

Dhillon, A., & Lockwood, B. (2004). When are plurality rule voting games dominance-solvable? *Games and Economic Behavior*, *46*, 55–75.

Elkind, E., Grandi, U., Rossi, F., & Slinko, A. (2015). Gibbard-Satterthwaite games. In *Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence, IJCAI 2015, Buenos Aires, Argentina, July 25–31, 2015* (pp. 533–539).

Elkind, E., Grandi, U., Rossi, F., & Slinko, A. (2017). Cognitive hierarchy and voting manipulation. arXiv preprint arXiv:1707.08598.

Farquharson, R. (1969). *Theory of voting*. New Haven: Yale University Press.

Feddersen, T. J., Sened, I., & Wright, S. G. (1990). Rational voting and candidate entry under plurality rule. *American Journal of Political Science*, *34*(4), 1005–1016.

Fishburn, P. C. (1978). A strategic analysis of nonranked voting systems. *SIAM Journal on Applied Mathematics*, *35*(3), 488–495.

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

Grandi, U., Hughes, D., Rossi, F., Slinko, A. (2017). Gibbard-Satterthwaite games for k-approval voting rules. CoRR, arXiv:abs/1707.05619.

Kawai, K., & Watanabe, Y. (2013). Inferring strategic voting. *American Economic Review*, *103*, 624–662.

Moulin, H. (1979). Dominance solvable voting schemes. *Econometrica*, *47*, 1337–1351.

Moulin, H. (1981). Prudence versus sophistication in voting strategy. *Journal of Economic theory*, *24*(3), 398–412.

Myerson, R., & Weber, R. (1993). A theory of voting equilibria. *American Political Science Review*, *87*(1), 102–114.

Obraztsova, S., Markakis, E., & Thompson, D. R. M. (2013). Plurality voting with truth-biased agents. In *Proceedings of the 6th International Symposium on Algorithmic Game Theory*.

Pattanaik, P. K. (1973). On the stability of sincere voting situations. *Journal of Economic Theory*, *6*(6), 558–574.

Satterthwaite, M. A. (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.

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

Sinopoli, F. D., & Iannantuoni, G. (2005). On the generic strategic stability of Nash equilibria if voting is costly. *Economic Theory*, *25*(2), 477–486.

Slinko, A., & White, S. (2008). Non-dictatorial social choice rules are safely manipulable? In *Proceedings of the 2nd International Workshop on Computational Social Choice (COMSOC-2008)*.

Slinko, A., & White, S. (2014). Is it ever safe to vote strategically? *Social Choice and Welfare*, *43*, 403–427.

Thompson, D. R. M., Lev, O., Leyton-Brown, K., & Rosenschein, J. S. (2013). Empirical analysis of plurality election equilibria. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS-2013)*.

Xia, L., & Conitzer, V. (2010). Stackelberg voting games: Computational aspects and paradoxes. In *Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-2010)*.